|CURIEUX|

ACADEMIC REVIEW

October Issue

Issue 12 Part 1

Editing Staff

Chief-In-Editor

George Hasnah

Lead Editor

Amanuel Geremew

Ashley Du

Assisting Editors

Jing Yan Huang

Nidhi Potineni

Sophia Kong

Tony Yuan

Ridit Aggarwal

Manu Sharma

Rishi Gadikota

Eric Wang

Table Of Contents

Page 6: Seismic Energy Harvesting Revised Research Paper By Lohitaksh Gupta

Page 13: Opioids and Non-opioids in Pain Management for Surgical Patients By Zhixiao Liu

Page 22: What Emptiness Reveals: Digital Archaeology in the Abandoned Villages of Aragon By Guillermo Srba Cabeza

Page 33: The Impact of AI Tools on Employees' Operational Efficiency and Decision Making By Peter Hung

Page 42: AI and Dental Caries: of Detection, Prediction, and Clinical Implementation By Jason Noh

Page 62: Strength vs. Unity: A Comparative Analysis of Nationalist Rhetoric in Trump and Xi Jinping's Foreign Policy Speeches By David Meyers

Page 77: A Structural Analysis: How Incentives Shape the Value-Action Gap of Climate Change By Nikki Samios

Page 93: Assessing the Role of In Vitro Fertilization in South Korea's Demographic Crisis By Gloria Kim

Page 98: Extreme Heat Is a Public Health Emergency in Massachusetts By Deetya Prasad

Page 101: Energy Aware Terrain Mapping and Autonomous Recovery Landing of a UAV on A Moving Ground Dock Station By Almaz Sagnalin

Page 111: The Effect of Dynamic Message Signs on Traffic Accidents By Mudit Gupta

Page 128: Explaining Differences in Human Cognition: A Biological, Cognitive, and Visual Framework By Akshita Tiwari

Page 139: Acceptable Risk, Unacceptable Harm: The Deficiency of Federal Regulation and the Exacerbation of Slow Violence in Cancer Alley By Krish Kocharlakota

Page 151: How to Score? Data Analytics Based Strategies for Varsity Level Volleyball By Koken Sugita

Page 160: Old Tennis Shoes: How Steinbeck's Simplicity Has Shaped My Journey as a Reader and Member of a Community By Tyler Temerowski

Page 164: Democratizing Failure Analysis: A Natural Language Processing Approach to Identifying Technical and Operational Bottlenecks in Early-Stage Micro-SaaS Ventures By Vaibhav Narayanan

Page 172: The Influence of Personality on Trust By Nichapa Khoncham

Page 179: Investigation Into How the Social Pressure Arising From Social Media's Content Virality Presents Promotional Benefits and Drawbacks for Indie Musicians By Alexis Roges

Page 193: A Comparative Study of Youth Political Agency and the Socio-Cultural Barriers to Legislative Power in Norway and South Korea By Myeounguk Kim

Page 204: Optimizing Growth Media for *Acinetobacter Baumannii* using Genome-Scale Metabolic Modeling By Jaden de Venecia

Page 223: Functional Applications Of CRISPR/Cas9 In Plant Biology: A Review Of Molecular-To-Phenotypic Translation By Mai Chi Tran

Page 246: Who Benefits from AI? Operational Efficiency and Cost Pass-Through in Quick Commerce Platforms By Tanishkka Aggarwal

Page 269: How Accurate is Your Diagnosis? The Effects of COVID-19 on Mental Illnesses & Diagnoses By Yunn Nway Lin

Page 276: Coupled Multi-Body Dynamics and Aerodynamic Load Modeling for Predicting Limit Handling in High-Downforce Hypercars By Aanya Garg

Page 295: Comparing Narrative Recall After AI Summarization and Full-Text Reading By Rohan Murthy

Page 317: System Design: Modern Large Language Models for Speech To Braille and Braille To Speech By Om Sandip Patel

Page 332: AI-Driven Marketing's Impact - Unpacking Measurable and Intangible Effects in B2B and B2C Contexts By Aarez Osman

Page 344: Co-inoculation of AMF and *Trichoderma* on Tomato and Basil Production By Alicia Naggah

Page 354: When to Spray: A Rainfall-Lagged Optimal Control Model for Desert Locust Intervention Timing in Pakistan By Mustafa Salman

Page 375: The Psychological Impact of School Bullying on Children and Adolescents in China By Yunhan Chen

Page 383: Assessing the Influence of ESG Considerations on Early-Stage Venture Capital Decision-Making By Rishav Mehta

Page 402: Why the Separation of Church and State is Necessary for Democracy to Continue By Violet Day

Page 408: Gold Mining and Its Effects on Cocoa Production in Ghana By Leo Cheng

Page 425: Evaluating Cultural and Linguistic Bias in Large Language Models: An Indian Context Study By Sayesha Dod

Page 443: From Concern to Choice: Climate Change and Sustainable Clothing Adoption By Arhan Harchandani

Page 455: Unraveling Democratic Backsliding: A Literature Review of Competing Actor-Centered, Structural, and Institutional Explanations for Democratic Backsliding By Logan Steelberg By Logan Steelberg

Page 467: Understanding, Treating, and Preventing Red Meat Allergy By Gary Peng, Uistean Travers, Siri Vanam

Page 474: Estimating the Progression of Normal Tension Glaucoma Using AI by Predicting Visual Fields By Brody Taniguchi

Page 482: The Impact of Acute Stress on Working Memory and Decision Making Performance in Young Adults in Kazakhstani Population By Mari Zhurba

Page 491: Evaluating the Optimal Characteristics for the Addition of Basalt Fibers in Metakaolin-based Geopolymer Materials By Ethan Yao

Page 508: A Theoretical Framework for the Use of Vedic Mathematics in Children with Learning Disabilities and Hearing Loss: Potential Cognitive Mechanisms and Future Research Directions By Aariana Gawande

Page 523: The Volatility Spillover Between Commodity Prices and Stock Market Indices By Zacharia Mir

Page 540: Gain-Loss Framing and Loss Aversion in Financial Decision Making By Pranay Doppalapudi

Page 549: Behavioral and Lifestyle Predictors of Obesity: A Quantitative Analysis of Dietary Habits, Physical Activity, and Technology Use Using Classification Modeling By Valerie Wu

Page 564: Juvenile Recidivism and Incarceration: A Comparison of U.S. Juvenile Justice Systems By Ava Miltenberg

Page 575: Why Do Young Consumers Spend Thousands of Dollars Traveling for Concerts? By Zexin Liu

Page 587: Is an Unenforced Law Really a Law? By David Liang

Page 600: Investigation on Temperature-dependent Mechanical Properties of Several Typical Polymers By Kexin Deng[1], Jiaheng Li[2], Derek W. Davies[3] and Junhui Hu

Page 611: From Plato's Cave to Algorithmic Feeds: Idealization and Moral Development in the Age of Social Media By Srihari Padmanabhan

Page 627: A Multi-Criteria Evaluation of Nuclear Reactor Technologies for Kazakhstan's First Nuclear Power Plant By Alikhan Zhetpis and Adel Abzhan

Page 641: Deep-Sea Mining: Assessing Environmental Threats and Buoyant Flocculation Rover Solutions By Aaron Zhang, Chris Kim, Noah Kim, Jayden Hwang, Daniel Kim and Ian Oak

Page 656: The Science Behind the Trends: Evaluating Popular Acne Remedies on Social Media Against Clinical Evidence By Aashna Gupta

Page 677: Why Nations Fail Review By Yağmur Eren

Seismic Energy Harvesting Revised Research Paper By Lohitaksh Gupta

1.1 ABSTRACT

This research aims to calculate the energy that can be generated by the energy harvesting of **seismic waves** from earthquakes in a region over a year and therefore analyse the efficiency of the process to conclude if it can be used as a clean energy source while working as a backup energy system for these areas. This is to predict if such harvesting could be used in **earthquake-prone** cities, the research has been done with an innovative plan of the **grid system**- pyramidal structures across streets with energy storage and connection to each other- of the device to show the scope of the seismic energy. The energy generated bya single mechanism in the device is estimated to be 5.8W, which is also generated by 3 ther mechanisms in different directions in the device, multiplying the output by 3, with 50% efficiency. The system that has been used has this mechanism 3-4 sides, generates more energy, the muti-axis capture may increase duty factor, but has to be validated experimentally. The aim of the study is to predict if such harvesting of energy is possible in the future, to run cities, and develop more technology on this. The findings of this study have been taken from various sources, to calculate the efficiency and the energy. The is inspired from other studies and the system of the device has been thought of to make this more realistic and available to maximise the use of the device.

1.2 INTRODUCTION

With major renewable energy sources such as hydro, wind and solar, research into clean energy is expanding as humans are trying to replace non-renewable energy completely. While there has been limited research into seismic energy harvesting because of its unpredictable nature and doubt in its efficiency, there is great potential that could be used. The few research papers that have been published explore the efficiency of a mechanism that can achieve this, and theoretically, it is possible to have better efficiency of the system. The earthquakes have long acted as destructors that cause great damage to the infrastructure and the power grids, the damage to the buildings is caused by the strong vibrations from the movement of tectonic plates which is because of the energy release from the core, making the city lose its power and functioning. If used as a power source, the energy from these movements could be converted into useful electric energy instead of mechanical waves, mitigating the effect of the earthquakes.

A mechanism to harvest this energy would work like the power grids to power the cities at the time of the earthquakes which would allow the city to work efficiently.

The research into the topic is scarce with different mechanisms proposed, based on single direction waves and a single device. Such as the one mentioned in a paper (1), which proposes a mechanism that captures waves only in one direction and only predicts the energy with a different application of the system. Their research works with microseisms and the idea of the system was different as well. Moreover, a paper published (2), refers to the dual purpose of seismic energy harvesting, which gives a lot of information into this topic.

In contrast and hope for higher efficiency, the mechanism described in this paper aims to use the

waves from all directions, and looking at the possibility of the earthquakes' waves spreading out in all directions, the mechanisms act as a part of the power grid which can help use a fraction of the large amount of energy released by an earthquake. The protection of the structure is done by its strong tetrahedral design.

2 METHODOLOGY

2.1 Procedure

This works on the basis of electromagnetic induction, the seismic energy released by earthquakes is converted into electricity. The vibration of the ground has been calculated according to the equations provided by the sources. To prove and investigate the efficiency of the mechanism, a case study is conducted on Budta, a place in the Philippines, that is around 3.5 magnitude. The energy reaching the mechanism with the peak ground acceleration and the frequency is calculated, with the efficiency of the mechanism assumed to be 50%. The structure of the mechanism and the planning has been described, assuming the distance and the number of mechanisms to predict if this would be a good backup system.

2.2 BACKGROUND INFORMATION

2.2.1 EARTHQUAKES

The earthquakes as traditionally known are a great source of energy, especially mechanical energy that is transferred through the movement of rocks. The mechanical energy, as elastic rebound theory says, can be stored in rocks (3). The short movement in the two ends of the faults elastically strains the rock near the fault. When the elastic strain exceeds the frictional strength of the fault, a hypocentre forms at the weakest point. From the hypocentre, the rupture rapidly propagates along the surface of the fault, causing the sides of the fault to slip past each other. In the opposite reaction to the force exerted by the rocks' frictional stress, rocks along the fault release energy. The energy in rocks is in the form of elastic energy, which is released in the form of heat. The series of waves radiate from the hypocentre in all directions, producing earthquakes. The amplitude of the seismic waves can range from micrometers (millionths of a meter) to tens of meters. The amount by which the seismic waves deform the ground decreases with distance from the earthquake.

2.2.2 MEASURING EARTHQUAKE

The most popular kind of magnitude is the richer magnitude to measure the strength of the earthquake, the other one is moment magnitude which is derived from analysis of all waveforms recorded from the shaking.

Moment(MO)= rigidity x area x slip

where rigidity is the strength of the rock along the fault, area is the area of the fault that slipped, and slip is the distance the fault moved. Thus, stronger rock material, or a larger area, or more

movement in an earthquake will all contribute to producing a larger magnitude (4).
Then Moment Magnitude (MW) :

$$M_{w} = \frac{2}{3}\log_{10}(M_0) - 9.1$$

Where:

● *MO* is the Moment.

Another way to measure the magnitude of the earthquake is by measuring its energy released, energy magnitude. The moment magnitude could be converted into energy magnitude by the following equation:

$$logE = 5.24 + 1.44Mw$$

Format the equations

2.2.3 SEISMIC WAVES

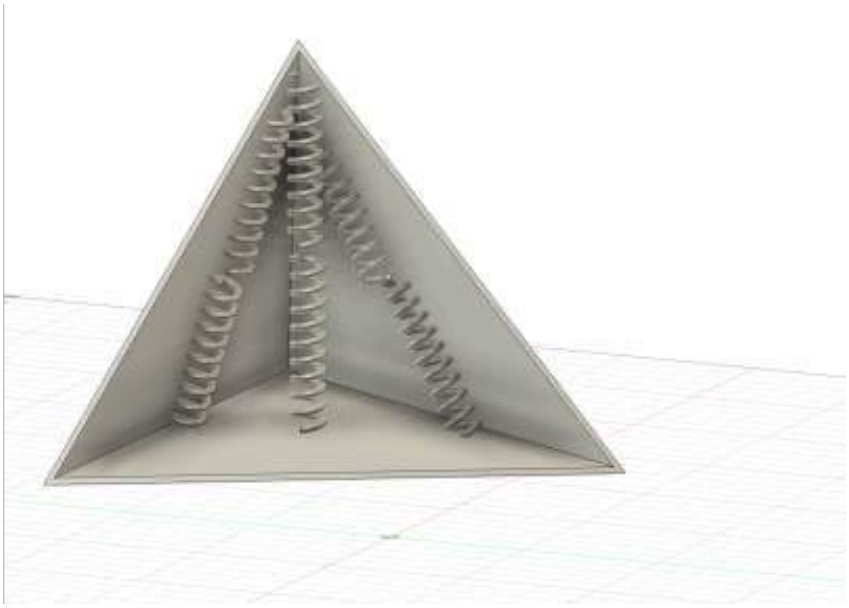
The wave as it travels from the epicentre, attenuates, which means that it wears down or the energy is lost from the wave as it is disrupted (5).

This equation can be rearranged to calculate the energy that reaches the mechanism. However there are other ways to calculate the power or energy at the point from the earthquakes, mentioned as pinpoint power.

3 MECHANISM

3.1 THE STRUCTURE

The mechanism has been designed in a tetrahedral structure to maintain its stability during earthquakes. Inspired by mechanisms used in previous papers published, incorporating a design that converts the waves from all directions, it offers more efficiency. The three faces of the pyramid and the vertical space in the middle, all are occupied by 2 springs each on two ends and magnet, directed by the set path carved in the face. The two springs help change the direction of the momentum of the magnet, with coils surrounding each of them. The constant change in momentum and the velocity of the magnet help with the current generation. In comparison to the mechanisms mentioned in other sources, the probability of the device capturing the wave is increased to the maximum. Similar to the one being investigated in the paper, a research explores one mechanism, which could be the base of all the calculations of energy output from the device.



3.2 THE EFFICIENCY

The efficiency of the mechanism is highly dependent on the selection of the springs and the material used to make it. A similar study on harvesting using Belleville springs was done, which estimates the equation for output power with specifics of the mechanism. These particular specs of the mechanism could be changed to make it more efficient. The mechanics, however, allow it to respond effectively to low frequency, broadband vibrations typical of seismic sources. The system converts kinetic energy into electrical energy via permanent magnets oscillating inside coils. The prototype described in the paper achieved an operational bandwidth from approximately 120-180 H, producing power up to several milliwatts depending on the number of magnets and input acceleration.

The equation for the input power from the seismic excitation can be estimated using the following equation:

$$P_{\text{input}} = {}^{1}_{2}\, m\, A^2 \omega^3$$

Where:

- m: effective mass of the moving components (kg),
- A: peak ground acceleration (m/s²),
- ω: angular frequency of excitation (rad/s).

To estimate A and ω from an earthquake of magnitude M, you can use the following empirical approximations:

1.Peak Ground Acceleration (PGA):

$\log_{10}(PGA) = a + bM - \log_{10}(R + c)$

For typical rock site conditions, a=−1.5, b=0.5, c=10

Where:

● R is the distance from the epicentre in kilometres
● M is the richer-scale magnitude.
Convert the result to linear units with A=10^log10(PGA).

2.Dominant Frequency:
f≈10/M殲ω=2πf
P =m . damping ratio . resonant frequency. maximum relative displacement

4 CASE STUDY

4.1 FREQUENCY OF EARTHQUAKES IN BUDTA, PHILIPPINES

In 2024, Budta was ranked number 7 for number of earthquakes according to the earthquake list [8], with 290 earthquakes yearly. The average magnitude of the earthquakes is 3.5. In earthquakes of this range, a developing country like Budta, has electrical failures that need to have backup systems. The mechanism is to be installed in a system like electric poles and the grid, which could produce continuous supplies of electricity during earthquakes and store it, to be used as a backup system for lights.

4.2 CALCULATION OF ENERGY FROM EARTHQUAKES

E= 5.24+1.44M

Magnitude M=3.5
Which is the average magnitude of the earthquakes in Budta, Philippines.
Distance R=1 km

As explained previously, the system will be like that of the electric grid or the water hydrants, which will be spread around the city, mainly around the expected epicenters. This will decrease the distance between the epicenter and mechanism to a few hundred metres.

PGA A ≈ 0.2116 m/s^2
Dominant frequency f ≈ 2.857 Hz (ω ≈ 17.94 rad/s)
Using these values of PGA and Dominant Frequency:
Input mechanical power Pin≈11.517

Output electrical power at 50% efficiency Pout ≈ 5.7585 W
The mechanism has the springs and magnets facing all the directions that would produce more energy and the mechanism placed together as a grid would be more efficient.

5 ANALYSIS AND CONCLUSION

The power output is enough to power emergency lights. If 80 percent of the number of electric poles in Budta are set up then they could be used together to power the emergency lights. Though the power generated by the mechanism is quite low, the theoretical value is the minimum that can be generated through the mechanism as the specs have been kept to a

minimum. This value can be multiplied by 3 or 4 to count the different spring-magnet systems in the mechanism, this will produce 15 W at least, that can power an emergency light for 2-3 hours. The device could be used as a back up system for emergency light in developing or underdeveloped countries that are prone to earthquakes and do not have rigid electrical systems which can sustain earthquakes. This with storage will also store the electrical energy from microseisms that do not affect the energy grids. The power produced is 5.785 W by one mechanism in a device, the grid could be used to make the seismic energy consumed better and mitigate the effect of the mechanical waves. The waves that cause the vibrations in the buildings are being used in the system to generate electricity which will dampen the waves upon the buildings.

If researched more and experimented on more, then the mechanism could be improved to be more efficient, and used as a proper electrical system that uses energy from microseisms and earthquakes. Scientists and researchers could implement a similar structure as suggested in this paper, in a small earthquake-prone zone to validate these results. They could further make changes to the features of the mechanism that would increase the efficiency and the grid- system, to make this feasible in larger areas and cities. The implementation of such an energy system could change the face of renewable energy, bringing an innovative approach to the use of seismic energy, dampening the effect of the earthquakes on cities' building and energy systems, and providing electrical energy.

Works Cited

1. Díaz-Mora, F., & González-Fallas, A. (2022). Energy harvesting from seismic waves for electricity production. E3S Web of Conferences, 336, Article 00035. https://doi.org/10.1051/e3sconf/202233600035 \ https://www.e3s-conferences.org/articles/e3sconf/pdf/2022/03/e3sconf_icegc2022_00035.pdf
2. Loong, C. N., Dimitrakopoulos, E. G., & Chang, C.-C. (2023). Nonlinear electromagnetic energy harvester–structure system under seismic excitation: Vibration mitigation and energy scavenging. Journal of Engineering Mechanics, 149(9), Article 04023058. https://doi.org/10.1061/JENMDT.EMENG-7107 orcid.orgascelibrary.org
3. Boore, D. M. (1977, December). The motion of the ground in earthquakes. Scientific American, 237(6), 68–76. https://www.jstor.org/stable/24953873
4. U.S. Geological Survey. (n.d.). Earthquake magnitude, energy release, and shaking intensity. Retrieved September 8, 2025, from https://www.usgs.gov/programs/earthquake-hazards/earthquake-magnitude-energy-release-and-shaking-intensity
5. Banerjee, S., & Abhishek, K. (2016). Determination of seismic wave attenuation: A review [PDF]. Disaster Advances, 9(6). Indian Institute of Technology Guwahati. https://www.iitg.ac.in/abhiak/wp-content/uploads/2019/05/Coda-wave normalization.pdf
6. Florindo, F., Acocella, V., Carlton, A., D'Odorico, P., Duan, Q., Gettelman, A., Halekas, J., Harris, R., Mollenhauer, G., Robock, A., Stirling, C., & Yokoyama, Y. (2023). 60 years and beyond of Reviews of Geophysics.Reviews of Geophysics. Advance online publication. https://doi.org/10.1029/2023RG000807 usgs.gov
7. Castagnetti, D. (2015). A Belleville-spring-based electromagnetic energy harvester. Smart Materials and Structures, 24(9), 094009. https://www.researchgate.net/publication/282929561_A_Belleville-spring-based_electromagnetic_energy_harvester
8. EarthquakeList.org. (2025, September 7). Global earthquake reports and recent news. https://earthquakelist.org/

Opioids and Non-opioids in Pain Management for Surgical Patients By Zhixiao Liu

Abstract

Opioids have traditionally been used in pain management for surgical patients due to their speed and effectiveness in managing pain, but a shift towards researching and implementing alternative methods of pain management using non-opioid medications has begun in recent years. The risk of dependency and negative side effects from using opioids has placed an emphasis on non-opioid medications such as ketamine, gabapentin, and nerve blocks. Two major approaches that have been used are the Enhanced Recovery After Surgery(ERAS) Protocol, which emphasizes preoperative and postoperative care to lessen damage during surgery and optimize recovery, and Opioid-Free Anesthesia (OFA), which replaces the use of opioids during surgery with combinations of non-opioids such as ketamine and local anesthetics. These strategies have been associated with fewer side effects, shorter hospital stays, and overall better recoveries, though limited data and challenges in implementation remain as barriers to widespread adoption of these strategies.

Introduction

Although opioids are one of the strongest and most effective options for pain management after surgery, perioperative prescribing practices can significantly increase the risk for opioid dependency and overdose in the coming months and years. One important part of surgical care and also medicine as a whole, pain management, was not established as a field of medicine until the 1960s (Collier 2018). While opioids have been traditionally used for postoperative pain management, there has been an increasing focus on finding alternative options, especially after the growing opioid crisis in the US and further research into their addictive properties (Phillips et al. 2017).

Opioids work by activating receptors on nerve cells. Normally, calcium ions can pass through channels in the nerve cells, which stimulates the release of excitatory neurotransmitters. These neurotransmitters cross the synaptic cleft (Dhaliwal & Gupta 2023) and bind to receptors, stimulating the pain response. On the other hand, when opioids are introduced into the bloodstream, they bind to the opioid receptors on the nerve cells, which then block the calcium channels, reducing the amount of calcium ions that can pass through and alleviating the pain as a result. Opioids also cause the neurons that release dopamine to fire more frequently. This causes euphoria, or a feeling of intense happiness, which is one of the main reasons why there is such a high risk of dependency. Understanding the effects of opioids, especially in the case of side effects and dependency, clarifies the growing emphasis on alternative, non-opioid pain management strategies and proves that it is necessary for better patient outcomes.

There are many advantages to using opioids for postoperative patients, some of which include their speed and effectiveness in treating pain, as well as their potency. Opioids are very powerful analgesics that provide very quick and significant relief from pain, especially when it is caused by surgery, injuries, or trauma. This especially has benefits for postoperative patients,

where pain control is imperative. An example of this is fentanyl, as it has a quick onset but short duration, making it useful for trauma codes (Häske 2017). However, there are also many disadvantages to using opioids, some of which can be very dangerous, including opioid dependence and overdose. When a person uses a large amount of opioids for a long period of time, they may become dependent on them. This is because over time, the brain begins to generate more opioid receptors and decreases the natural production of endorphins. Opioids can also disrupt the brain's natural dopaminergic system, diminishing the pleasure a person can experience from normal everyday activities and becoming reliant on opioids for dopamine. Opioid dependence can also increase the risk of overdose as a person begins to take higher and higher doses to achieve the same level of analgesia. An opioid overdose occurs when the amount of opioids overwhelms the respiratory receptors in the body, leading to respiratory depression and death if untreated (Paul et al. 2021). Postoperative patients have a higher risk of opioid dependence, especially if they have already used opioids before, leading to the emphasis on finding alternative, non-opioid options to treat those patients.

Drug Type	Mechanism of Action	Side Effects	Duration of Effect
Opioids	Bind to GPCR receptors in the brain and spinal cord → blocks calcium channels and prevents neurotransmitters from firing	Dizziness, nausea, respiratory depression, risk of dependence	**Short-acting:** 2–6 hrs **Long-acting:** 12–24 hrs
Ketamine	Blocks NMDA receptors in the brain → reduces pain signal transmissions and induces dissociative state	Hallucinations, changes in sensory processing, risk of dependence	**IV**: 15-30 minutes **Oral**: 1-2 hours
Gabapentin	Blocks transmission of pain signals in brain and spinal cord similarly to opioids and ketamine	Respiratory depression, altered mental status, behavioral changes	5-7 hours
Local Anesthetics	Blocks sodium-gated channels, prevents nociceptors from firing→affected nerves contained to one area	Infection, bleeding at injection site, nerve damage	30 minutes-12 hours depending on location and specific anesthetic used
Tylenol (Acetaminophen)	Block COX enzymes→prevents production of prostaglandins and thromboxanes	Nausea, allergic reaction, risk of liver damage from overdose	4–6 hours
NSAIDs	Block COX enzymes similarly to Tylenol	Stomach upset, heartburn, risk of kidney damage from overdose	6-8 hours

Table 1: Characteristics of Drug Types Used For Surgical Patients

Several different alternatives to opioids have been suggested for the management of postoperative pain, including nerve blocks, gabapentin, ketamine, Tylenol, and NSAIDS. A nerve block involves the injection of a local anesthetic near nerves to block the action potential, which conveys pain signals to the central nervous system, including the brain. The transmission of pain is prevented by blocking sodium-gated channels at nociceptors, the anesthetic acting as a barrier to prevent nerves from sending pain messages to the spinal cord and brain. This does not affect the rest of the body, allowing patients to regain mobility sooner and allowing faster recovery overall. This procedure can even replace general anesthesia in selected cases, as shown for the Wide Awake Local Anesthesia No Tourniquet (WALANT) technique. However, the performance of nerve blocks requires trained specialists and may also be associated with infection, bleeding, and nerve damage complications. Another alternative to opioids is ketamine, a dissociative anesthetic that works by blocking NMDA receptors, reducing the perception of

pain (Kurdi et al. 2014). Compared to opioids, ketamine maintains airway reflexes and acts as a bronchodilator. This makes it particularly useful for patients with respiratory problems. Although ketamine has some disadvantages, like psychological dependency and potential cardiovascular complications, it very rarely leads to lethal overdose and is followed by mild withdrawal symptoms compared to opioids (Liu et al. 2016).

Other non-opioid analgesics include gabapentin and acetaminophen/nonsteroidal anti-inflammatory drugs. Gabapentin, an anticonvulsant with calcium channel-blocking properties, is effective in decreasing postoperative and neuropathic pain while reducing the need for opioids or other medications to be used alongside it (Peng et al. 2007). However, in combination with opioids, it may increase the risk of respiratory depression and altered mental status, causing the FDA to issue warnings against using gabapentin for people with underlying respiratory disease or risk factors. Acetaminophen and NSAIDs reduce pain through the inhibition of enzymes responsible for the production of inflammatory mediators (Ghlichloo & Gerriets 2023). Their use in combination produces central and peripheral analgesia and reduces opioid intake (Jaeschke et al. 2020). However, excessive dosing can result in severe organ damage, including liver failure with acetaminophen and renal injury with NSAIDs.

Although all these options for postoperative pain management have many benefits, they can also have negative effects. One of the biggest reasons for finding an alternative to opioids is how addictive they are. Many of the most commonly prescribed opioids are also most commonly involved in deaths from overdose. Patients who are prescribed opioids after a surgery are 44% more likely to use opioids long-term, and 14.7% of patients report using opioids without a prescription, longer than prescribed, or in greater amounts than prescribed (Hah et al. 2017). Over 60% of patients who are prescribed 90 days of continuous opioid therapy continue to use it years later. Opioids can also cause severe side effects when administered, including dizziness and nausea as well as respiratory depression, where the opioids can block neurotransmitters in the brain that control breathing (Kosten & George 2002). Patients who are given opioids lose the respiratory drive to breathe, and this is a leading cause of death from overdose. While non-opioid options for pain management seem promising and have had some positive results, there has still not been enough testing done, causing medical professionals to be reluctant to start the shift towards non-opioids due to the possibility of side effects that have not been discovered.

Due to the addictive properties of opioids and its potentially dangerous side effects, it is important for possible alternatives to be studied. Recently, multimodal pain management, the use of multiple medications from different drug classes in order to use no or as few opioids as possible, has become more popular. One example of this are the Enhanced Recovery After Surgery protocols, or ERAS. This system is designed to achieve a quicker and easier recovery after surgery by maintaining organ function and decreasing the body's stress response caused by surgery. It focuses on prevention and control of pain and nausea, less intervention from doctors, and eliminating opioid use. Another example is Opioid-Free Anesthesia (OFA), where instead of using opioids during surgery, a combination of non-opioid medications is used. This method also

limits the amount of side effects a patient experiences, including nausea or respiratory depression, which are both associated with opioids.

ERAS Protocol

The Enhanced Recovery After Surgery protocol, or ERAS protocol, is a new protocol for surgery patients that emphasizes patient education and perioperative care. It is one of the main alternatives to traditional practices using opioids. The ERAS protocol includes all phases of an operation, including the preoperative, intraoperative, postoperative, and post-discharge phases. This protocol has been shown to decrease the risk of complications, improvements in cardiopulmonary function, and a quicker return to normal activities (Melnyk et al.). The main principles of the ERAS protocol include preoperative counselling and education, optimization of nutrition, standardized anesthetic regimens, and early mobilization.

Unlike other, more traditional surgical practices, the ERAS protocol includes and puts an emphasis on the preoperative phase of a surgery, or the days and hours before it. One large part of this that the ERAS protocol specifically focuses on is nutrition. This protocol especially addresses malnutrition, which is a preoperative risk factor that can lead to poor surgical outcomes. Proper nutrition can also decrease the risk of complications and reduce the length of hospital stays, as well as decreasing the risk of surgical mortality. The ERAS protocol suggests preoperative carbohydrate loading through the use of clear carbohydrate beverages up to 2 hours before surgery. This is a part of the ERAS protocol that differs from traditional surgical practices, where most patients are advised to fast after midnight the day of the surgery to reduce the risk of aspiration. The ERAS Protocol argues that preoperative fasting can actually increase metabolic stress, hyperglycemia, and insulin resistance.

ERAS also advises that patients stop all smoking and alcohol intake 4 weeks before surgery, with the help of their primary care physician if needed. The guidelines also recommend patients stop oral contraceptives and hormone replacement before surgery to decrease the risk of perioperative venous thromboembolism, or VTE, a serious surgical complication where blood clots form in the veins and restrict blood flow to important organs (Altman et al.). Some patients are also given medications, such as antiemetics, NSAIDS, and antibiotics, in order to lessen postoperative side effects and decrease chance of infection.

One factor that ERAS guidelines take into account during the intraoperative phase of surgery is body temperature. Hypothermia, where the body temperature falls below 36 Celsius or 96.8 Fahrenheit, is a common occurrence in surgical patients, with around a 20% to 30% chance. This can significantly increase the risk of complications such as surgical site infections, prolonged wound healing, and increased blood loss. ERAS uses multiple strategies to prevent hypothermia, including warmed intravenous fluids, warming blankets, and increasing the surgical room temperature. ERAS also focuses on maintaining euvolemia, or optimal fluid balance. Hypovolemia, a reduction in blood volume, is common in surgical patients where there is significant trauma or blood loss. ERAS guidelines encourage intake of clear fluids in the hours before the surgery and fluid therapy.

Arguably the most important phase of ERAS is the postoperative phase. Nutrition is also important in this phase, as it can help maintain muscle strength and speed up healing. Guidelines recommend returning to a regular diet within 24 hours, which can promote earlier bowel function and greater patient satisfaction. Chewing gum is also advised, as it can reduce the incidence of postoperative ileus, or a temporary slowing or stopping of bowel movement after surgery. Maintaining a patient's blood glucose level to around 180 to 200 mg/dL, as perioperative hyperglycemia is associated with poor surgical outcomes, is also recommended.

Early mobilization is another crucial aspect of the ERAS protocols. Early mobilization helps reduce muscle atrophy, compromised pulmonary function, thromboembolic complications, and overall results in shorter hospital stays. The guidelines recommend mobilizing patients, such as sitting, standing, and walking, within 24 hours of surgery and for multiple hours a day. The ERAS protocol also minimizes opioid use, instead using medications such as NSAIDS and local anesthetics, with opioids only used for breakthrough pain if necessary.

There are many benefits to implementing ERAS protocols, as it can reduce the stress on the body from surgery and help the patient return to normal quicker. It includes the stages before a surgery as well as during and after, which can help patients prepare more. It also educates patients more on the surgery, helping increase patient satisfaction. By incorporating more local anesthetics and non-opioid pain management, it also reduces the amount and severity of side effects and prevents opioid dependence. There is also a lower rate of complications due to preventative measures such as early mobilization and antibiotics before the surgery. This also reduces costs for the hospital from shorter stays and low complication rates.

The ERAS protocol also has specific medication routines and guidelines for different types of surgery, with medications given before, during, and after the procedure. In the preoperative phase, medications such as antiemetics or prophylactic antibiotics may be given to reduce the amount and intensity of side effects and reduce the risk of infection. In the intraoperative stage, regional anesthesia is given if necessary and a combination of non-opioid medications are given to reduce the need for opioids and anesthesia. Tranexamic acid, a medication that prevents bleeding by inhibiting the breakdown of blood clots, can also be given to reduce the risk of bleeding during or after surgery (Engelman et al. 2019).

Other than medication, there are also many other things that can be done to prepare for surgery. A hemoglobin test may be performed, as low hemoglobin can increase the risk of complications such as delayed wound healing or increased blood loss. For patients with nutritional deficiencies, oral supplementation can be started 7 to 10 days before the procedure as well as continuing to have clear liquids up to 2 hours before surgery. This can lead to better outcomes for patients and also reduce pain, decreasing the need for opioids.

The ERAS Protocol has been found to significantly reduce the risk of side effects and optimize patient outcomes when compared to simply using opioids. One study found that the average length of the hospital stay after a surgery decreased by 25% when following the ERAS Protocol. There was a 72% reduction in overall opioid use, and 16% of patients were completely opioid-free for the first 3 days after their surgery (Meyer et al. 2019).

OFA

Opioid-free anesthesia, or OFA, is another alternative for traditional pain management practices. Opioids have been the most commonly used medication for postoperative pain, however there can be severe complications that come with using them. Efforts to find alternatives to using opioids in medical settings have greatly increased since the start of the opioid crisis in the 1990s in the United States. The greatest risk that comes with using opioids is respiratory depression, which can also increase the risk of aspiration.

Opioid-free anesthesia focuses more on the intraoperative stage of a surgery. Unlike traditional practices, opioids are not used during the surgery and only as needed after. In OFA, small amounts of multiple non-opioid medications are used to block multiple neurotransmitters. The usage of different medications may also play a role in reducing the amount and severity of side effects. For example, ketamine can be used to provide pain relief but also to prevent side effects such as tachycardia or hypertension (Sakan et al. 2023). OFA can also prevent side effects that may occur from a large dose of a specific medication through using small doses of several. Local anesthetic, such as lidocaine, is also used and is proven to reduce the need for opioids and reduce inflammation at the surgical site.

The protocols for OFA include beginning medications up to an hour before surgery. Medications such as gabapentin, midazolam, and dexmedetomidine are given preoperatively to reduce anxiety or stress and for their sedative effect. Low doses of ketamine and lidocaine are also given close to intubation. A mixture of the above medications as well as muscle relaxant or magnesium if needed, is given to maintain the anesthesia during the surgery. The anesthetic is usually stopped around 20 minutes before the patient awakens. Antiemetics, NSAIDS, and local anesthetic at the surgical site are all administered depending on the patient's state. OFA is associated with lower postoperative opioid consumption, a decrease in intubation time if needed, and shorter ICU stays.

Discussion

While opioids have been proven to be efficient in providing pain relief for postoperative patients, the risks associated with them have led to research into alternative options. These options include ketamine, which is potent and fast-acting similar to opioids but can also keep patients calm. A downside to ketamine is the risk of dependence, also similar to opioids, however the withdrawal symptoms are less extreme. Another option are nerve blocks, which numb a specific part of the body, allowing for better mobility and a quicker return to normal function, although it must be performed by a medical professional, forcing patients to remain in the hospital for longer, and has the risks of infection or nerve damage. A third alternative is gabapentin, which has a similar mechanism of action to opioids and ketamine and can reduce the amount needed of other medications. However, it can cause respiratory issues, as well as altered mental status. The fourth option are Tylenol or NSAIDS, which are widely available and easily used, but have a low potency and a risk of overdose. Two new protocols for postoperative patients that minimize the use of opioids are the ERAS protocol and OFA. Both protocols,

especially the ERAS protocol, focus on not only the intraoperative stage but also the pre- and postoperative stages. Different medications are given to ensure that the least possible amount of opioids are used, reducing hospital stays, side effects, and increasing patient satisfaction.

Similar to other surgical practices, there are also some disadvantages to the ERAS protocol. A main disadvantage is the complexity of the protocols, which requires coordination from the patient as well as multiple professionals, which can be challenging for hospital teams to implement. The individualized or hospital-specific treatment plans may also make it difficult to assess the overall effectiveness and compare to other practices. A main factor that makes it difficult to implement not only the ERAS protocol but also other new practices is the reluctance of medical professionals, especially those who are older or in hospitals who already have established practices, to learn and implement new ones.

The advantages of OFA include reduced opioid-related side effects such as nausea, respiratory depression, and postoperative ileus, or disruption of the bowels. It also enhances recovery similar to ERAS protocols, leading to greater patient satisfaction and shorter hospital stays. It also reduces the risk of opioid dependence by minimizing or eliminating opioid use, addressing the ongoing opioid crisis. OFA also especially has benefits for specific patients, such as those who are obese, have respiratory conditions, or patients with a history of opioid use. The disadvantages of OFA are similar to those of the ERAS protocol, including a need for specialized training and teamwork of hospital staff and difficulty standardizing the treatment plan.

Alternatives to opioids in not only postoperative care but also other aspects of medicine are important and should be a priority. There are many advantages to the different medications and protocols listed above that can improve the overall experience of patients and also be beneficial for the hospital. For example, shorter hospital stays and reduced side effects mean that the patient can return to normal life quicker, and also uses up less resources from the hospital. It can also reduce the risk of opioid dependence and lead to better surgical outcomes. However, these alternatives have not undergone enough testing to be widely adopted by hospitals. The process to incorporate opioid-free medications should be gradual, as there is not much evidence yet of the effectiveness of these medications in the long term and in larger groups of people. With more studies and research, the use of non-opioids in surgery as well as in medicine overall can become more proven and widespread. The next steps for future studies can include figuring out the effects of non-opioids in patients long-term as well as the effects for more types of surgeries and groups of people. Currently, as the shift to non-opioids is still relatively recent, there is still limited data on the effects on non-opioid protocols in patients months or years after surgery. The data that does exist is still limited to more common and less invasive surgeries.

While opioids are a very useful tool in pain management for surgical patients, there are other non-opioid medications that can be just as useful with fewer downsides, such as local anesthetic, ketamine, gabapentin. The ERAS protocol and OFA are two protocols that have been created to shift the focus in surgical procedures from opioids to other medications. While they both have shown benefits, there is further research that needs to be done before the change from opioids to non-opioids becomes widespread.

Works Cited

Collier, Roger. “A Short History of Pain Management.” Canadian Medical Association Journal, vol. 190, no. 1, 7 Jan. 2018, pp. E26–E27, www.ncbi.nlm.nih.gov/pmc/articles/PMC5760261/,https://doi.org/10.1503/cmaj.109-5523.

Phillips, Jonathan K., et al. Pain Management and the Intersection of Pain and Opioid Use Disorder. Www.ncbi.nlm.nih.gov, National Academies Press (US), 2017, www.ncbi.nlm.nih.gov/books/NBK458655/.

Dhaliwal, Armaan, and Mohit Gupta. “Physiology, Opioid Receptor.” PubMed, StatPearls Publishing, 24 July 2023, www.ncbi.nlm.nih.gov/books/NBK546642/.

“Prescription Opioids.” CAMH, 2018, www.camh.ca/en/health-info/mental-illness-and-addiction-index/prescription-opioids.

Paul, Alok K., et al. “Opioid Analgesia and Opioid-Induced Adverse Effects: A Review.” Pharmaceuticals, vol. 14, no. 11, 27 Oct. 2021, p. 1091, www.ncbi.nlm.nih.gov/pmc/articles/PMC8620360/, https://doi.org/10.3390/ph14111091.

Kurdi, Madhuri S, et al. “Ketamine: Current Applications in Anesthesia, Pain, and Critical Care.” Anesthesia: Essays and Researches, vol. 8, no. 3, 2014, p. 283, www.ncbi.nlm.nih.gov/pmc/articles/PMC4258981/, https://doi.org/10.4103/0259-1162.143110

Peng, Philip WH, et al. “Use of Gabapentin for Perioperative Pain Control – a Meta-Analysis.” Pain Research and Management, vol. 12, no. 2, 2007, pp. 85–92, https://doi.org/10.1155/2007/840572.

Häske, David, et al. “Analgesia in Patients with Trauma in Emergency Medicine.” Deutsches Aerzteblatt Online, vol. 114, no. 46, 17 Nov. 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5730701/, https://doi.org/10.3238/arztebl.2017.0785.

Ghlichloo, Ida, and Valerie Gerriets. “Nonsteroidal Anti-Inflammatory Drugs (NSAIDs).” PubMed, StatPearls Publishing, 1 May 2023, www.ncbi.nlm.nih.gov/books/NBK547742/.

Jaeschke, Hartmut, et al. “Novel Therapeutic Approaches against Acetaminophen-Induced Liver Injury and Acute Liver Failure.” Toxicological Sciences, 11 Jan. 2020, https://doi.org/10.1093/toxsci/kfaa002.

Lucas, Guillherme Nobre Cavalcanti, et al. “Pathophysiological Aspects of Nephropathy Caused by Non-Steroidal Anti-Inflammatory Drugs.” Brazilian Journal of Nephrology, vol. 41, no. 1, 2019, pp. 124–130, www.ncbi.nlm.nih.gov/pmc/articles/PMC6534025/

Hah, Jennifer, et al. “Chronic Opioid Use after Surgery: Implications for Perioperative Management in the Face of the Opioid Epidemic.” Anesth Analg, vol. 125, no. 5, Nov. 2017, pp. 1733–1740, www.ncbi.nlm.nih.gov/pmc/articles/PMC6119469/pdf/nihms-1504139.pdf

Kosten, Thomas, and Tony George. “The Neurobiology of Opioid Dependence: Implications for Treatment.” Science & Practice Perspectives, vol. 1, no. 1, July 2002, pp. 13–20, pmc.ncbi.nlm.nih.gov/articles/PMC2851054/, https://doi.org/10.1151/spp021113.

Melnyk, Megan, et al. “Enhanced Recovery after Surgery (ERAS) Protocols: Time to Change Practice?” Canadian Urological Association Journal = Journal de l’Association Des Urologues Du Canada, U.S. National Library of Medicine, Oct. 2011, pmc.ncbi.nlm.nih.gov/articles/PMC3202008/#:~:text=3-,Enhanced%20recovery%20after%20surgery%20(ERAS)%20protocols%20are%20multimodal%20perioperative%20care,profound%20stress%20response%20following%20surgery.

Altman, Alon D., et al. “Enhanced Recovery after Surgery: Implementing a New Standard of Surgical Care.” CMAJ : Canadian Medical Association Journal, vol. 191, no. 17, 29 Apr. 2019, pp. E469–E475, www.ncbi.nlm.nih.gov/pmc/articles/PMC6488471/, https://doi.org/10.1503/cmaj.180635.

Engelman, Daniel T., et al. “Guidelines for Perioperative Care in Cardiac Surgery.” JAMA Surgery, vol. 154, no. 8, 1 Aug. 2019, p. 755, jamanetwork.com/journals/jamasurgery/fullarticle/2732511

Meyer, Larissa A., et al. “Effect of an Enhanced Recovery after Surgery Program on Opioid Use and Patient-Reported Outcomes.” Obstetrics and Gynecology, vol. 132, no. 2, 1 Aug. 2018, pp. 281–290, www.ncbi.nlm.nih.gov/pmc/articles/PMC6245646/#:~:text=Overall%2C%20patients%20in%20the%20ERAS, https://doi.org/10.1097/AOG.0000000000002735.

Sanja Sakan. “Opioid Free General Anesthesia in Clinical Practice – a Review Article.” Acta Clinica Croatica, 1 Jan. 2023, https://doi.org/10.20471/acc.2023.62.02.15.

What Emptiness Reveals: Digital Archaeology in the Abandoned Villages of Aragon
By Guillermo Srba Cabeza

Abstract

Throughout the twentieth century, hundreds of villages in the Spanish autonomous community of Aragon were lost to rural exodus, hydraulic engineering projects, and decades of economic neglect. But what initially presents as a demographic catastrophe has actually generated an unintended opportunity. Huge tracts of formerly inhabited landscape now lie open, unbuilt, and increasingly overgrown with vegetation—ideal conditions for archaeological exploration using remote-sensing technology.

The central argument to this paper puts forth that Aragon's depopulation has created a uniquely readable archaeological environment. Through technologies including LiDAR (Light Detection and Ranging), GIS (Geographic Information Systems), and drone-mounted photogrammetry, layers of Iberian, Roman, Visigothic, and Islamic settlement that dense modern occupation would have obscured or destroyed are now being revealed. This paper draws on landscape archaeology, remote sensing research, and official heritage data from the Sistema de Información del Patrimonio Cultural Aragonés (SIPCA) to explore the history of Aragonese depopulation through case studies from the Àger and Espique Valleys. It demonstrates that where people have disappeared, technology has stepped in to generate archaeological knowledge at a pace and scale previously impossible. Further, the role of digital media and visualization tools–dimensions of the work as culturally significant as the discoveries themselves–is explored in how these findings are ultimately translated into publicly accessible resources.

Introduction

My family comes from the tiny Teruel village of Berge, situated in the rural southern portion of Aragon. Like most villages in that part of Spain, Berge has been losing people for over a century. The landscape around it–scrubland, limestone ridges, dry riverbeds–holds the remains of a Visigothic burial ground, a strand of the layered occupation that characterizes nearly every square kilometer of Aragon. Few people outside the village know it exists and fewer still have studied it systematically. The burial ground itself rests in a landscape that is, in the starkest demographic sense, abandoned.

That abandonment, and what it makes possible, is the foundation of this paper. Aragon is one of the most sparsely populated regions in Western Europe. Its three provinces–Huesca, Zaragoza, and Teruel–cover roughly 47,000 square kilometers (an area larger than Denmark), yet Teruel province alone has a population density of less than ten inhabitants per square kilometer (Instituto Nacional de Estadística, 2025). This is not a natural condition but rather, the product of a century of migration to Spain’s cities, hydraulic projects, and political neglect that emptied villages at a rate that scholars have described as “extreme” even by the standards of rural southern Europe (Collantes and Pinilla, 2004). In an earlier study, it was estimated that two

hundred villages in the province of Huesca alone have been completely or partially abandoned (Acín Fanlo, 1997). Across all of Aragon, the figure is higher still.

In many ways, this wholesale emptying of the villages of Aragon can be seen as a tragedy, and in human terms it is. But, there is a resulting phenomenon that offers the promise of a sort of redemption: depopulated landscapes, precisely because they lack the soil disturbance, construction activities, and legal complexity of inhabited ones, are extraordinarily well-suited to remote-sensing archaeology. In short, when people leave, the land stops being modified. Ruins stabilize rather than being eroded, demolished, and replaced. Vegetation grows back over archaeological features, preserving their outlines rather than leaving them exposed to the corrosive elements of time and weather. And when a technology arrives that can strip away that vegetation digitally–rendering the bare earth beneath in three dimensions, at one centimeter resolution, all from the air–the results can be remarkable.

The technology that accomplishes this feat is called LiDAR. Combined with Geographic Information Systems (GIS) used for spatial analysis, drone-mounted cameras for photogrammetric documentation, and digital visualization platforms for freely disseminating this information to the public, it is transforming what archaeologists can now know and study about landscapes that human occupation once obscured. Throughout this paper, it is argued that Aragon's depopulation has created an accidentally singular archaeological opportunity, that remote-sensing technologies are beginning to capitalize on this moment, and that digital media has a crucial role to play in making those discoveries meaningful far beyond academic settings.

A Region Emptied: The Making of an Abandoned Landscape

The depopulation of rural Aragon did not begin with the Spanish dictator Francisco Franco, but his regime accelerated it dramatically. Researchers at the University of Zaragoza trace the roots of the collapse of Aragon’s rural economy to the nineteenth century, when the absence of fertile land for seasonal sheep migration, repeated crop failures of subsistence agriculture, and the loss of cottage textile work to factories began pushing people toward cities (Collantes and Pinilla, 2004). The mountainous areas of Huesca, where terrain made farming difficult and transport and trade with urban centers nearly impossible, were hardest hit. By the early twentieth century, depopulation was already well underway.

Francoism exacerbated the scale and the coercive dimension of the process. The regime's hydraulic policy–the construction of hundreds of dams and reservoirs across Spain as tools for agricultural modernization and national propaganda–displaced tens of thousands of rural inhabitants. More than five hundred villages were submerged beneath reservoir waters, from Galicia in the extreme northwest of Spain to the Aragonese Pyrenees (Ecooo, 2024). This said, it is important to note that dam-building was not purely an invention of the Franco regime. The majority of Spain is prone to droughts and irregular flooding and efforts to manage water sources had begun in the last decade of the nineteenth century. The Franco era put a new spin on it, however, in the manner in which the hydraulic projects were executed. New projects were implemented without compensation adequate for the disruption and displacement they caused,

without democratic consultation, and under a political system that offered no legal recourse to those displaced (Lanero, cited in Newtral, 2025).

Jánovas, a village in the Pyrenean valley of the Río Ara in Huesca province, exemplifies the human cost with particular clarity. In 1951, the Franco regime approved a hydroelectric and irrigation reservoir project on the River Ara, a meandering waterway through the small villages of the Pyrenees. In anticipation of this project, families were systematically expropriated and forced out of their homes and towns during the 1960s, their houses left standing on the assumption that rising waters would soon erase them. In a cruel twist, however, the dam was never built. Fifty years after the project's approval, in 2001, an environmental impact assessment concluded the project was not viable, and it was formally cancelled. Properties began finally reverting to their former owners–or their descendants–only in 2010 (Aragondigital, 2026). For nearly half a century, Jánovas existed in a legal and physical limbo: emptied but not flooded, abandoned but not demolished, its streets and church and nineteenth-century suspension bridge preserved by the same inertia that had destroyed its community.

The story of Jánovas is extreme in its injustice, but the demographic pattern it represents is not. Across Aragon, villages emptied through combinations of economic marginalization, outward migration to industrial centers, failed infrastructure projects, and the continued collapse of traditional agriculture that accelerated through the 1960s and 1970s. Teruel province, where my family's village of Berge is located, is today the most depopulated province in Spain. Of Teruel's 236 municipalities, 218 have fewer than a thousand inhabitants, and the province as a whole registers a population density of just 9.19 inhabitants per square kilometer (Instituto Nacional de Estadística, 2025). That density has been falling, with brief interruptions, since 1910.

What this means archaeologically is the subject of the rest of this paper.

The Technology: LiDAR, GIS, and the PNOA Revolution

LiDAR (Light Detection and Ranging) has been the proverbial candle in the dark responsible for the recent strides forward in the archaeological study of Aragon and Spain. It works by firing millions of laser pulses from an aircraft or drone, measuring the time each takes to return, and then constructing a precise three-dimensional terrain model from the results. Its crucial advantage over conventional aerial photography is that the laser actually penetrates vegetation: laser pulses passing through gaps in forest or scrubland canopy are able to reach the ground beneath, allowing software to strip out elements of vegetation and generate a bare-earth Digital Terrain Model (DTM). Archaeologists then read the DTM for the subtle topographical signatures of buried or surface remains–wall foundations, building outlines, field boundaries, irrigation channels–that are invisible to both the naked eye and to satellite imagery alike. GIS overlays that terrain data with historical maps, documentary sources, and any existing archaeological inventories, and builds composite pictures of how a particular landscape has been used across time. Drone-mounted photogrammetry operates at the site level, generating detailed three-dimensional reconstructions of individual structures. These serve both as research records

and as the raw material for digital heritage visualization–the 3D models, virtual site visits, and documentary video footage that translate these discoveries into public knowledge (Jouav, 2025).

Spain's national LiDAR dataset is called the Plan Nacional de Ortofotografía Aérea (PNOA), a systematic programme of aerial survey and LiDAR coverage commissioned to produce standardized, high-resolution terrain data across the whole country. When the Instituto Geográfico Nacional made it freely available in December 2015, it overnight put centimeter-resolution data for all of Spain–including Aragon's mountains, abandoned valleys, and depopulated villages–into the hands of every archaeological researcher without the need to commission a single dedicated flight (Monterroso-Checa, 2017). A 2020 assessment in the Boletín del Museo Arqueológico Nacional found rapid uptake by archaeologists across multiple historical periods. Particularly successful were the detections of Roman military installations, prehistoric megaliths, and medieval settlements in terrain that ground survey alone had never been able to reach (Cerrillo Cuenca and López López, 2020).

Impressive as it is, this dataset has real limits. Its standard resolution of 0.5 points per square meter is sufficient for the discovery of major structural remains but misses subtler features. Most recent archaeological work in Aragon therefore treats the PNOA as a good-quality first pass and then follows with targeted, higher-density drone LiDAR surveys where the evidence requires a more detailed assessment.

All of these technologies perform best in conditions that depopulated landscapes provide. Modern construction methods erase surface features, disturb historical archaeological layers, and introduce materials that distort remote-sensing signals. Vegetation regrowth does just the opposite: it stabilizes soil above archaeological features and remains penetrable by laser. The empty Aragonese countryside–terraced hillsides returning to scrub, village streets undisturbed for decades–is not an obstacle to remote-sensing archaeology. Rather, it is the ideal environment.

What the Silence Reveals: Case Studies in Remote-Sensing Discovery

The clearest demonstration of the investigative power of LiDAR in an emptied Spanish landscape comes from work done in the Espique Valley in the province of Granada. A team from the Biocultural Archaeology Laboratory (MEMOLab) at the University of Granada conducted a combined aerial photogrammetry and drone LiDAR survey of a mountain valley that has seen a steady loss of inhabitants over the past 150 years. The results, published in the journal Archaeological Prospection in 2025, were profound. Nineteen drone flights initially generated a point cloud of over 4.2 billion measurements. After LiDAR processing software stripped out the vegetation canopy to expose the bare-earth Digital Terrain Model, analysts identified sixty-seven locations of potential archaeological interest. Field inspection followed, confirming thirty-six previously unknown sites of archaeological interest. Among others, the team found entire buried building complexes, mining entrances, furnace depressions, and ancient irrigation channels (acequias) that historical records had never documented (Rodríguez Bulnes, Román Punzón, and Martín Civantos, 2025).

The Espique Valley is not Aragon, but its landscape closely mirrors that of Aragonese mountain territory. They both share a complex topography, heavy vegetation regrowth over formerly inhabited land, a layered multi-period archaeological record ranging from late Roman through Islamic occupation, and a depopulation history that stripped away the human activity that would otherwise have obscured or destroyed surface evidence. The MEMOLab team explicitly noted that the valley's archaeological heritage–both habitation sites and a range of productive spaces including agricultural terracing and water management infrastructure–had remained effectively invisible to prior survey precisely because the terrain was inaccessible and densely vegetated (Rodríguez Bulnes et al., 2025). LiDAR made it legible there as it is beginning to do in comparable landscapes throughout Aragon.

In Aragon itself, the Àger Valley straddles the border between the city of Lleida and the Huesca Pyrenees, a mountain territory that formed part of the medieval Aragonese frontier. This site provides a documented case of LiDAR and radar-based landscape archaeology applied to exactly the type of medieval rural settlement that characterizes much of the region. A doctoral study from the University of Lleida used LiDAR, radar, documentary sources, and select archaeological digs to reconstruct the valley's medieval settlement dynamics from the fifth through the thirteenth centuries. The project involved analysis of the spatial organization of fortified hilltops, churches, and agrarian parcel systems that defined almost a millennium of successive phases of occupation and abandonment in the region (Porcheddu, 2018). The study conclusively demonstrated that LiDAR-derived terrain analysis, combined with historical mapping and documentary evidence, could recover a detailed picture of historical settlement evolution in a landscape where ground-level visibility was severely limited by vegetation and terrain.

The findings of the Àger study matter just as much as how the archaeological remnants were found. LiDAR and radar data were key to revealing the spatial logic of a medieval frontier settlement system: hilltop fortifications positioned to control valley access routes, parish churches sited to serve clusters of dispersed farmsteads, and agricultural parcel boundaries fossilized in the current field pattern despite centuries of abandonment and reoccupation. These are not features that conventional ground survey could reconstruct across an entire valley in a single study — the vegetation cover and terrain complexity would make a comprehensive investigation on foot practically impossible. Remote sensing exposed the skeleton of a settlement landscape that historical documents could hint at but never fully describe. The significance for further investigation in Aragon is directly comparable. The Àger Valley's medieval archaeology belongs to the same cultural and historical world as the dozens of empty and overgrown frontier valleys in the Huesca Pyrenees that have yet to receive equivalent systematic surveys.

The broader Aragonese archaeological record that these technologies are beginning to illuminate is exceptional in its depth and variety. Aragon has been continuously occupied for thousands of years. Its territory encompasses Iberian and Celtiberian settlements, Roman colonial cities and road networks, Visigothic burial grounds and churches, and one of the most significant concentrations of Islamic-period artefacts in the Iberian Peninsula. The province of Teruel

alone–the most depopulated–contains an expansively impressive archaeological record. In just this one province exist such diverse findings as UNESCO World Heritage rock art sites in the Parques Culturales of Albarracín, Maestrazgo, and Río Martín; the Roman city of La Caridad at Caminreal; and the remarkable twenty-five kilometer Roman aqueduct connecting Albarracín, Gea de Albarracín, and Cella (Turismo de Teruel, 2021). The protohistoric settlement at Alfambra adds another layer still, with 2023 and 2024 excavation campaigns combining surface prospection, drone survey, and LiDAR to identify occupation spanning the Bronze Age through the sixth century BCE (VI CAPA, 2025).

Within this layered record, Visigothic archaeology holds a particular relevance to the landscape of depopulation. The Visigothic period, roughly the fifth through early eighth centuries CE, left a fragmented pattern of rural settlement and burial grounds across the Iberian interior, much of it in territories that subsequently underwent repeated cycles of abandonment and resettlement through the Islamic conquest, the Christian Reconquista, and the post-medieval rural exodus. Even Berge, the empty little town of my family, carries a documented Visigothic burial ground. It and countless other such sites scattered across the region remain poorly understood or researched, in part because with the emptying of the villages surrounding them there were no local community members to drive further research. The SIPCA database catalogues Visigothic archaeological evidence across all three Aragonese provinces, including maps of necropolises and funerary remains compiled by the Atlas de Historia de Aragón. But cataloguing is not the same as understanding. LiDAR and GIS-based survey of the territories around depopulated villages like Berge could potentially reveal the spatial organization of Visigothic settlements in the Bajo Aragón with a resolution and completeness that conventional fieldwork–requiring access, funding, and local knowledge that depopulation has removed–cannot achieve.

Why Emptiness Enables Discovery: The Archaeology of Depopulated Space

The relationship between human absence and archaeological visibility initially seems counterintuitive. Abandoned places are typically seen as sites where knowledge is lost: no one to remember, no one to maintain, no one to care. The oral histories of Aragonese villages, the vernacular knowledge of field names and water rights and seasonal migration routes, disappear with the people who held them. That loss is irreversible.

But the physical record of human occupation behaves differently from its human custodians. Inhabited landscapes are constantly being modified. Buildings are demolished and replaced, fields are ploughed to depths that destroy stratified deposits, roads are laid on top of ancient trackways, and the legal and logistical complexity of working in populated areas constrains where and how archaeologists can dig. Indeed, researchers using remote-sensing technologies have noted that some of the most pressing obstacles to LiDAR-based archaeological survey in Spain are not technical but structural–the presence of modern infrastructure, private ownership, and the "noise" introduced by contemporary land use (Cerrillo Cuenca and López López, 2020). Depopulation serves as the agent that removes these obstacles.

At the same time, the particular ecology of post-depopulation landscapes in Aragon–dry, scrubby, Mediterranean vegetation gradually reclaiming agricultural terraces and building plots–creates conditions that from a LiDAR perspective, are close to optimal. The region's particular type of sparse vegetation allows laser pulses to reach the ground while its presence stabilizes soil over features. Abandoned terraced hillsides retain their forms in the absence of continued cultivation, preserving the highly localized signatures of walls, building foundations, and field boundaries. The Espique Valley study found ancient irrigation channels–features of extraordinary delicacy, easily destroyed by a single season of mechanical ploughing–intact and detectable precisely because the valley had been depopulated long enough for vegetation to stabilize but not so long that erosion had erased the relevant topographic relief (Rodríguez Bulnes et al., 2025). Aragon's valleys exist in a very similar window of time, and it is not clear how long that window will remain open. Climate change is altering vegetation patterns; erosion is ongoing; the occasional attempts to repopulate or develop abandoned villages introduce the same types of disturbances that depopulation temporarily removed.

There is also an economic and ethical argument for prioritizing remote-sensing surveys in these Aragonese landscapes now. Traditional archaeological excavation is expensive, slow, and irreversible. It yields detailed results, but only for the precise ground disturbed, and that context can never be undone. In contrast, LiDAR and GIS-based surveys are non-invasive, cover vast areas efficiently, and produce data that can guide targeted excavation of the precise sites most likely to yield significant results. For regions like Aragon, where funding for heritage work is limited and the archaeological landscape is enormous, remote sensing is not a luxury but a necessity. The Cerrillo Cuenca and López López assessment noted that for many sites in Spain, LiDAR-PNOA data will represent the only systematic survey ever conducted — not a supplement to fieldwork but its complete replacement (Cerrillo Cuenca and López López, 2020). In a landscape as vast and as empty as Aragon's, this might be seen as a cause for despair, but it is not. Rather, it is a recognition of what is actually possible and the unique opportunity for exploration and discovery.

That said, LiDAR has meaningful limitations, but recognizing this actually strengthens rather than undermines the argument for its use. The PNOA dataset's standard resolution–0.5 points per square meter across most of Spain, rising to 1 point per square meter in some areas–is sufficient for detecting large structural remains such as foundations, terracing systems, and road networks, but falls short for smaller or subtler features. Cerrillo Cuenca and López López specifically noted that this resolution is inadequate for a significant range of archaeological feature types whose detection would require substantially denser point clouds (Cerrillo Cuenca and López López, 2020). This is precisely why the Espique Valley study supplemented PNOA data with purpose-flown drone LiDAR at far higher density, thus achieving the resolution needed to detect not just buildings but furnace depressions and irrigation channel profiles. For Aragon, this means that the freely available PNOA data provides an invaluable first-pass survey capable of identifying major sites and prioritizing areas for further investigation, but that realizing the full archaeological potential of these landscapes will ultimately require targeted,

higher-resolution drone surveys of the most promising zones. The PNOA made the landscape legible for the first time; drone LiDAR makes it readable in full.

Conclusion: Finding the Story, Then Telling It

Archaeology has always been about recovery. Recovering the material traces of lives lived, decisions made, structures built and abandoned is at the very heart of archaeological discovery. What is new in Aragon, and throughout depopulated rural Spain, is that a technology now exists capable of conducting that recovery on a profound scale, non-invasively, over terrain that conventional fieldwork cannot systematically reach. LiDAR has shown, in the Espique Valley and in the Àger Valley and in dozens of sites across the country, that emptiness is not the same as absence. The ground beneath the scrub and the abandoned terraces is teeming with evidence of lives lived and histories told.

But discovery is not communication, and this is where the other half of the argument becomes important. The archaeological record of Aragon–its Iberian hillforts, its Roman cities, its Visigothic burial grounds, its Islamic irrigation systems, its medieval fortified villages–belongs to everyone with a claim on that history. It belongs to the descendants of the families who left, to the Spanish and international research community, and to anyone with an interest in understanding how a region as historically rich as Aragon came to be so empty. Making that record accessible on a broad scale requires the same digital tools that help discover it. Photogrammetric 3D models of excavated sites, drone footage of LiDAR-surveyed landscapes, interactive GIS maps that let users navigate the heritage record of a comarca or a valley are not just decorative additions to the research. They are precisely the means by which archaeological knowledge jumps out from the specialist literature and reaches the people for whom it is most meaningful.

There is also a more immediate institutional dimension to this process of recovery. In Aragon, a site of archaeological interest can not be legally protected until it is formally inventoried–and obviously, it can not be inventoried until it is found. Providentially, the SIPCA functions not only as a research database but as the administrative gateway to heritage protection under the Ley 3/1999 del Patrimonio Cultural Aragonés. Any site that enters SIPCA's census as “Bien Inventariado” or “Bien Catalogado” acquires legal protections against development, illegal excavation, and neglect that an undocumented site entirely lacks (Gobierno de Aragón, 2022). Remote-sensing survey therefore has a direct legal consequence in that each site identified through LiDAR analysis and subsequently verified in the field can move through SIPCA into legally protected status. In a region where abandoned villages are sometimes targeted as potential development sites and where illegal metal-detecting and artifact removal are documented problems, this is not simply a bureaucratic technicality. It is the essential difference between a site with real archaeological relevance that survives and one that does not. The pipeline from LiDAR detection to SIPCA registration to legal protection is, in this respect, as important as the discoveries themselves.

The DiGHER project at the University of Zaragoza, which has created virtual visits to rock art and archaeological sites across Aragon, points toward what this process looks like in practice. The project uses digital photography, photogrammetry, and 3D modeling to create digital records of sites that are physically remote, ecologically fragile, or legally restricted — sites that most people will never visit in person but can now encounter through a screen (Universidad de Zaragoza / SIPCA). The same approach could be applied to the large scale discoveries that LiDAR is making possible. For example, a virtual flyover of a LiDAR-derived terrain model showing the outlines of a buried Roman villa, overlaid with the historical record and the story of the village that once surrounded it, would constitute a genuinely new form of heritage communication.

For the descendants of Aragonese emigrants–and there are many of us, scattered across Spain, across Europe, across the Americas–this matters in a way that goes beyond academic interest. The villages our families left, or were forced to leave, are not simply lost. Their histories are recoverable. The ground still holds a waiting record of what the former residents carried away in memory. LiDAR is beginning to read it. The task now is to make sure that reading reaches the people for whom it has the most meaning. This will rely on both better science and the full range of digital and cinematic tools that make better storytelling possible in the twenty-first century.

Aragon's depopulation was a tragedy. The archaeological record it left behind, now recoverable through technology, is a gift.

Works Cited

Acín Fanlo, José Luis. Paisajes con memoria. Viaje a los pueblos deshabitados del Alto Aragón. Prames, 1997.

Aragondigital. "El pueblo fantasma de Aragón que guarda el último puente colgante del siglo XIX que queda en España." Aragón Digital, 7 Jan. 2026, www.aragondigital.es.

Cerrillo Cuenca, Enrique, and Antonio López López. "Evaluación y perspectivas del uso del LiDAR en la arqueología española." Boletín del Museo Arqueológico Nacional, vol. 39, 2020, pp. 221–238.

Collantes, Fernando, and Vicente Pinilla. "Extreme Depopulation in the Spanish Rural Mountain Areas: A Case Study of Aragon in the Nineteenth and Twentieth Centuries." Rural History, vol. 15, no. 2, 2004, pp. 149–166. Cambridge University Press.

Ecooo. "La memoria perdida de los embalses franquistas." Ecooo, 2024, ecooo.es/blog/mito-agua-franquista-politica-pantanos/.

Gobierno de Aragón. "Todos los bienes del patrimonio cultural de Aragón en un clic." Go Aragón, Apr. 2022, www.goaragon.es.

Instituto Nacional de Estadística (INE). "Padrón Municipal de Habitantes." INE, 2025, www.ine.es.

Jouav. "LiDAR in Archaeology: How is LiDAR Used in Archaeology?" Jouav, 7 May 2025, www.jouav.com/blog/lidar-archaeology.html.

Monterroso-Checa, Antonio. "Remote Sensing and Archaeology from Spanish LiDAR-PNOA: Identifying the Amphitheatre of the Roman City of Torreparedones (Córdoba-Andalucía-Spain)." Mediterranean Archaeology and Archaeometry, vol. 17, no. 2, 2017, pp. 71–85.

Newtral. "Esto antes de Franco ya pasaba: la política de construcción de embalses que el dictador 'heredó'." Newtral.es, 28 May 2025, www.newtral.es/embalses-franco/20250528/.

Porcheddu, Andrea. "LIDAR for Landscape Archaeology: Detecting Archaeological Features in the Àger Valley." Medieval Territories, edited by Flocel Sabaté and Jesús Brufal, Cambridge Scholars Publishing, 2018, pp. 215–240.

Rodríguez Bulnes, Jesús, Julio M. Román Punzón, and José M. Martín Civantos. "Remote Sensing Analysis and LiDAR Experimenting in the Espique Valley (La Peza, Granada, Spain)." Archaeological Prospection, vol. 32, no. 2, 2025, pp. 395–407.

Sistema de Información del Patrimonio Cultural Aragonés (SIPCA). Portal web del SIPCA. Gobierno de Aragón / Instituto de Estudios Altoaragoneses, www.sipca.es. Accessed 1 May 2025.

Turismo de Teruel. "Patrimonio Arqueológico." Teruel Versión Original, 2021, www.teruelversionoriginal.es/patrimonio-arqueologico/.

Universidad de Zaragoza / SIPCA. "El Forau de la Tuta / Abrigo de La Vacada." DiGHER (Digital Heritage), Departamento de Ciencias de la Antigüedad, cartografía patrimonio cultural.unizar.es. Accessed 1 May 2025.

VI Congreso de Arqueología Profesional Aragonesa (CAPA). "Resultados preliminares del análisis del asentamiento protohistórico mediante prospección superficial, vuelos de dron y LiDAR." Colegio Oficial de Doctores y Licenciados en Filosofía y Letras y en Ciencias de Aragón, 2025, cdl-aragon.es/capa/.

The Impact of AI Tools on Employees' Operational Efficiency and Decision Making

By Peter Hung

Abstract

The rapid development of artificial intelligence (AI) raises an important question of to what extent do AI assistants enhance employee efficiency in the workplace. Understanding the degree to which AI assistants improve productivity may offer deeper insights into how AI has transformed daily work practices. While prior studies have explored the overall effect of AI on the whole business level, fewer have examined the effect it has on small companies and with data from employees. This research aims to determine both the perceived increase in working efficiency and the potential problems associated with the use of AI assistants on originality. This study addresses two key questions. First, to what extent do AI assistants increase employees' working efficiencies. Second, the employees' perceptions of AI assistants' underlying risks. Data was collected during June of 2025 through a survey distributed to 71 employees from two small technology companies operating in Taiwan. The survey used Likert-scale items to measure perceived increases in efficiency and multiple-choice questions to identify whether employees believe there are any problems related to using AI assistants. The study found that while AI assistants generally enhance employees' work efficiency, concerns remain regarding fairness, credit attribution, workplace dynamics, and data privacy, with variations in perceptions influenced by age, frequency of use, and departmental roles.

Keywords AI assistants, working efficiency, data originality, Taiwan

Introduction

The last decade has seen the rapid development of artificial intelligence (AI) in many businesses. Businesses across different industries now use advanced analytics and machine learning to remain competitive (Schmitt 1). Previous studies have established that AI-powered applications have the capability to learn continuously, make predictions by analyzing data, and generate ideas for developing new products, unlock deeper insights, uncover hidden patterns, and predict future trends (Abdulaziz Aldoseri 1). For example, AI algorithms in the healthcare sector are used to predict disease outbreaks and personalize treatment plans based on patient data. In retail, machine learning models help companies optimize inventory and recommend products to customers based on purchasing behavior. The ability to abstract other meanings from a large amount of data allows businesses to make data-driven decisions more easily (Enholm et al. 1721).

By using the advantages AI provides, companies may unlock a world of unlimited possibilities, enhanced competitiveness, and greater success in today's rapidly evolving AI landscape (Abdulaziz Aldoseri 1). In this paper, AI assistants are defined as AI-based technologies designed to assist people with everyday tasks and decision-making, such as ChatGPT, Google Gemini, DeepSeek, and Microsoft Copilot. These assistants rely on underlying

technologies such as machine learning, which have transformed how companies operate, execute and share insights.

Despite the growing presence of AI in organizational settings, relatively little is known about how individual AI assistants affect company operations and most existing papers discuss the overall use of data analytic technology.

This paper seeks to address the gap by analyzing the impact of AI assistants on employee working efficiency and identifying the potential problems associated with its use.

Google Gemini, Chat GTP, and Microsoft Copilot are some of the many available AI assistants. It can assist with simple tasks such as answering questions, and more importantly, analyze large volumes of data. Specifically, employees in a company may find this helpful because in their daily work, there is often large amounts of sales data or data from research and development. According to an analysis by IoT Analytics (Fernandez), Google Gemini is used by approximately 15% of the companies, Chat GPT with 9%, and Microsoft Copilot with 39%.

The study focuses specifically on Google Gemini, Chat GTP, and Microsoft Copilot, used by the two target companies in this research—InTeligo and LEOTEK in Taiwan. Selecting an AI assistant that the employees routinely use will ensure access to a sufficiently large user base and allow the study to examine real-world usage rather than hypothetical adoption. Consequently, other AI assistants were excluded, not owing to any presumed inferiority, but to maintain analytical depth and contextual relevance within the company under study.

This paper aims to provide a new insight into how AI assistants may affect employee efficiencies. While this study provides new insights into the role of AI assistants in enhancing productivity, it does not seek to encompass the full range of AI tools available in the market. This paper tests two hypotheses:

1.AI assistants positively impact employee operational efficiency.

2.The use of AI assistants entail various levels of risks related to fairness (work-related opportunities, promotions, rewards/awards), originality (ethics), privacy (leak of company data), people-relationship (company culture), and accuracy.

Research Design

A participatory quantitative research approach is employed in this study. A quantitative approach was employed since the data generated through the survey is able to ensure that the responses are standardized, reducing bias from the participants, and to help identify the pattern.

The survey sample size consists of 71 participants. The participants of this study were recruited from the employees of various departments of two tech companies in Taiwan that use AI assistants in their daily work. Data from this study were collected through a survey that includes three likert scale questions and seven multiple choice questions. Each Likert scale questions were rated on a scale from 1 (no improvement) to 5 (significant improvement), the results were quantified in this way so that the employees perceived improvement of their working efficiency can be examined. The survey was sent out to the participants through emails

and all the responses were collected anonymously. The survey allowed a detailed understanding of participants perception and an in depth analysis of the percentage and trend of participants' opinions. Graphs and figures are presented in this study to illustrate the findings.

In the survey, the independent variables are the 4 commonly used functions related to AI assistants—data analysis, article translations or summarizations, email writing, and note taking. These functions were selected because they are the most frequently used functions for the day-to-day business operations. Therefore, analyzing the employees' perceptions of these functions could contribute to the understanding of the impact of AI assistants on working efficiency. Additionally, the survey also considers independent variables related to the potential risks of using AI assistants, specifically whether the use of such AI tools have a negative effect on fairness, data privacy, originality, and people's relationship in the company. The dependent variables in this survey focus on the employees' responses around the topics related to the extent people rely on AI assistants. These include the employees' frequency of using each function, the degree that each function improves employees' working efficiency, and their perception of potential problems that are caused by AI assistants.

Results

1. AI assistants and work-efficiency gains

The question asked 71 respondents how each AI-assistant function affected their work efficiency. Figure 1 illustrates the distribution of categories among the listed functions, representing main functions of AI assistants. What is interesting about the data in this table is that improvements differed markedly across functions. For instance, translating and summarizing articles (n=49, ≈ 69%) and writing emails (n=41, ≈ 58%) are functions that received a noticeably high number of people that think it significantly improved their working efficiency. However, for the functions analyzing data and taking notes, the number of respondents reporting a significant improvement was comparable to those indicating a moderate improvement, with approximately 24-29 participants.

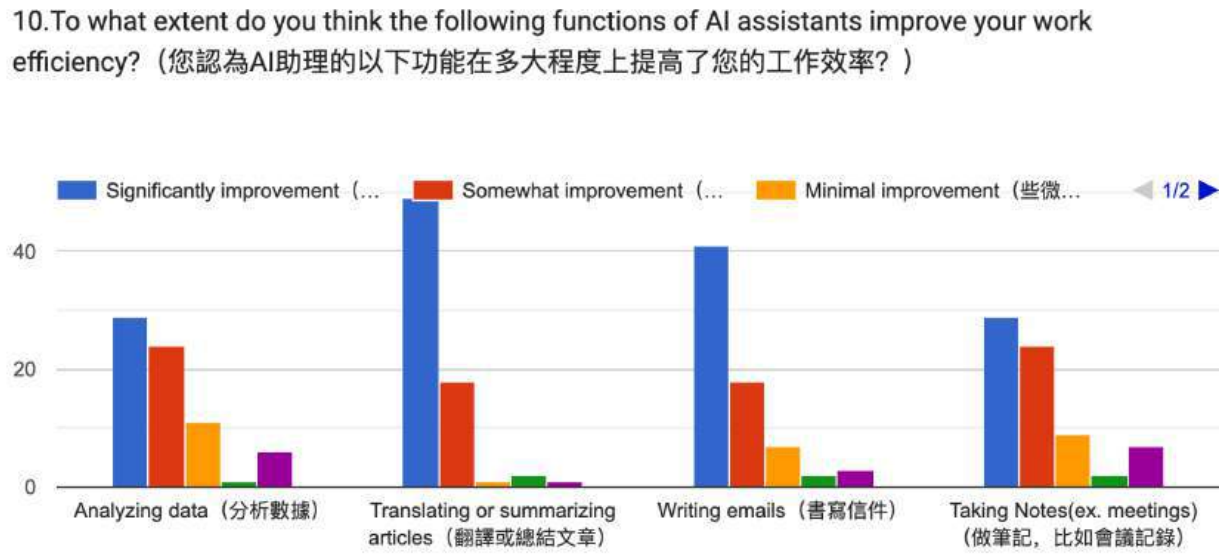


Figure 1: AI assistant and its impact on work efficiency

2. AI Risk Perceptions By the Frequency of AI Use

2.1 AI assistants' impacts on the fairness of the working environment

This set of data examines whether usage frequency is associated with perceived risks to workplace fairness. The survey shows that 23.9% of the 71 respondents think that AI assistants will affect the fairness of the working environment, while 76.1% do not. However, this result varies between the frequency they use AI assistants. According to Figure 2, for participants that are daily users (n=37), only 18.9% believe AI threatens fairness, whereas up to 34.7% of people that use AI a few days a week (n=23) think there is a risk, which is nearly twice the percentage. This shows a result that people who use AI more often are less likely to perceive an impact on fairness, while people who use AI less often think otherwise.

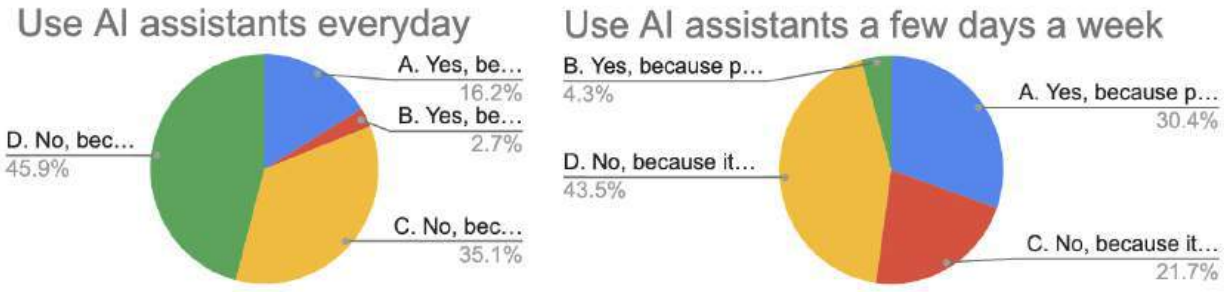


Figure 2: AI assistants' impacts on the fairness of the working environment

2.2 AI assistants and employee credit attribution issues

The question addresses if an AI assistant makes attribution harder depending on the frequency of employees' usage. The survey results indicate that over half of the 71 participants think that there would be no issue since the content is not directly used. However, according to Figure 3, among the participants who see potential challenges, the reasons differ based on the frequency they use AI assistants. 24.3% of daily users (n=37) worry that if multiple employees reach similar AI-generated ideas, credit cannot be assigned. On the other hand, people who do not use AI often (n=23) were more likely to perceive attribution as challenging, with 30.4% noting that data was often shared among co-workers, making it difficult to identify its original source.

Figure 3: AI assistants and employee credit attribution issues

3. AI Risk Perceptions by Departments

3.1 AI assistants' impact on work environment concerns

Figure 4 illustrates the relation between the different departments to which the participants belong and how they think AI assistants affect the working environment. According to the 100% stacked bar chart, we can see that the departments, R&D, Business, and Product-Management respondents (n=57) exhibited similar patterns, with > 60% selecting the disagree options (C and D). However, it is shown that none of the Marketing respondents (n=8) perceived a negative impact on the working environment.

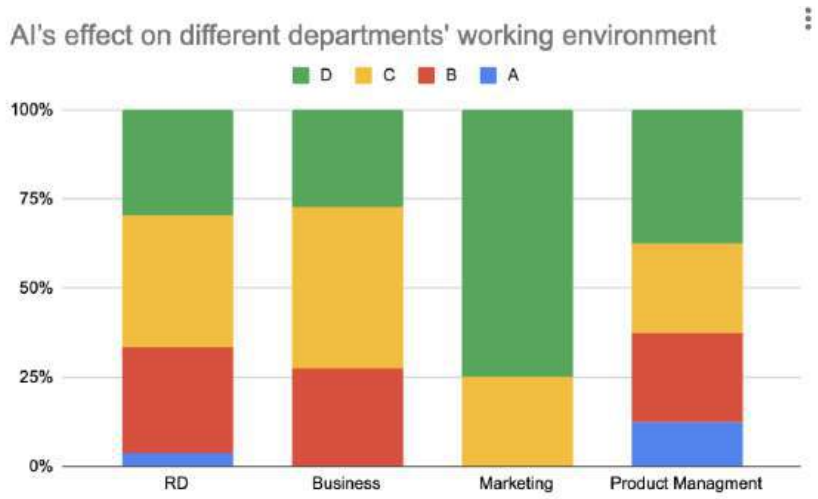


Figure 4: AI assistants' impact on work environment concerns

3.2 AI assistants' impact on data privacy concerns

We analysed whether perceived data-privacy risks differ by departments. According to Figure 5, the product management department (n=10) has a greater concern with data privacy, with 40% selecting agreeing options (C and D), while other departments have about 0-15%. Another notable finding is that the marketing department (n=8) has a higher percentage of people concerned about personal data privacy, whereas the business department expressed greater concern regarding company data privacy.

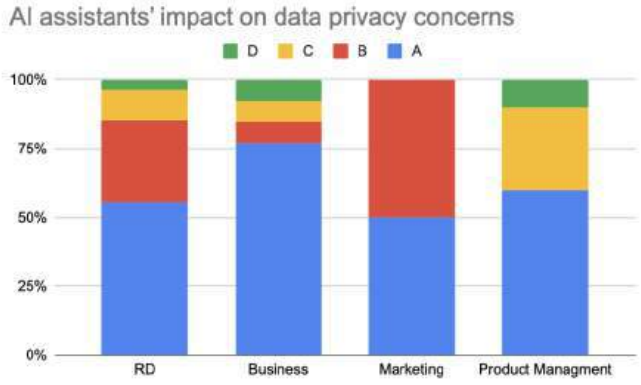


Figure 5: AI assistants' impact on data privacy concerns

4. AI Risk Perceptions by Age

4.1 AI assistants' impact on workplace fairness

The data below illustrates generational differences in perceptions of workplace fairness influenced by the use of AI assistants, particularly in areas such as job opportunities. Figure 6 showed that younger participants (n=12, 22-36 year-old) have a higher percentage (53%) that thinks AI does not affect workspace fairness since it's only used as a tool. In contrast, older

participants (n=52, 37-52 year-old) have a slightly higher percentage of people thinking there is a concern, with 16.7% in the 22–36 age group and 26.9% in the 37–54 age group, approximately 10% higher.

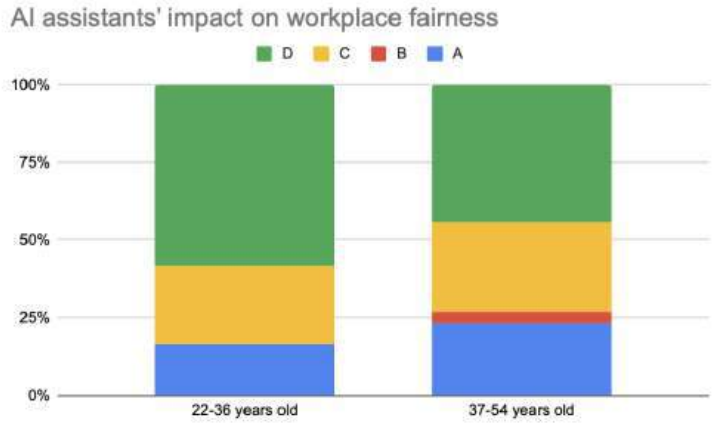


Figure 6: AI assistants' impact on workplace fairness

Discussion and Conclusion

This present study was conducted to examine the impact of AI assistants on the working efficiency of employees' daily work. The study set out to determine whether the use of AI contributes positively to the efficiency and identify underlying problems that exist for the implementation in the workplace. This research paper also analyzed the trends and relationships between the underlying problems and participants' departmental affiliation and their age groups.

The findings of this study indicates that the majority of the respondents perceive that the four primary functions of AI: analyzing data, translating or summarizing articles, writing emails, and taking notes, have an improvement in their daily work. Notably, the function translating or summarizing articles was most frequently reported as significant improvement, while other functions demonstrate an even distribution between "significant" and "somewhat" improvement. This may be attributed to the fact that translating and summarizing articles is more of a useful function for people nowadays, especially for working when you might need to read lengthy studies or news articles, learning unfamiliar materials.

An initial objective of the project was to identify the relationship between the participants' age and the frequency of AI use with underlying concerns related to fairness and credit attribution. As the results presented in section 2.1, 2.2, and 4.1, the most important finding was that more frequent and younger AI users expressed fewer concerns about fairness or credit attribution; whereas frequent and older users reported greater concerns about it. This pattern may be explained by frequent and younger users who are more skilled and confident with AI assistants, so they tend to view it as a supportive tool. On the other hand, less frequent and older users may possess limited skills or knowledge, making them more likely to see AI as a potential threat to opportunities or attributions.

The third question addressed the relationship between the employees' departmental affiliation and their perception of how AI affects the working environment. As indicated in

section 3.1, marketing staff reported fewer concerns about Al's negative impact on the working environment compared to other departments such as RD, business, and product management. A possible explanation for this is because marketing roles focus on external client interactions and sales, in which AI does more of supporting their work rather than competing eternally. However, other departments have more internal performance evaluation, which makes AI feel more like a competitive factor.

Another aim of this study was to determine the relationship between the employees' departmental affiliation and their perception of how AI might cause a threat to data privacy. The results show that marketing expressed greater concerns about customer data privacy, while business departments are more focused on the protection of company data. This distinction might be explained by the fact that marketing handles personal customer data frequently since they directly interact with clients, raising fears of leaks through AI tools. On the contrary, business departments handle sensitive operational and strategic company data, making them more focused on protecting internal business information.

These results suggest that AI is a significant part of employees' daily work practices. However, there still exist problems that should be considered and noticed as well, particularly about fairness, credit attribution, working environment, and data privacy. This study provides organisations with insights into these underlying concerns expressed by respondents, and the hidden trends that may be overlooked by companies. The study also provides possible explanations for these trends, so companies can develop strategies or policies to mitigate such risks. These findings contribute to a clearer understanding of the extent of AI's impact on the workspace, and also drawing attention to the potential risk of AI before they use them.

A limitation of this study is the relatively small sample size, the data were only collected from two small size tech companies in Taiwan, resulting in a total of 71 respondents. This constraint was due to the restrictions with obtaining organizational support, limited willingness to fill out the survey and the maximum number of potential participants within small companies. This limitation limits the possibility that some results may not accurately capture the patterns, and could result in overgeneralized trends. For future research, it would be valuable to aim for a larger sample size, a wider range of companies and larger organizations, allowing greater participants and more diverse perspectives. It would also be better if the data is collected from broader age groups and departments, so there is a more comprehensive understanding of trends and other hidden problems related to adopting AI assistants in the workplace.

Work Cited

Aldoseri, Abdulaziz, et al. "AI-Powered Innovation in Digital Transformation: Key Pillars and Industry Impact." Sustainability, vol. 16, no. 5, Feb. 2024, p. 1790, https://doi.org/10.3390/su16051790.

Alghamdi, Noorah A., and Heyam H. Al-Baity. "Augmented Analytics Driven by AI: A Digital Transformation beyond Business Intelligence." Sensors, vol. 22, no. 20, Oct. 2022, p. 8071, https://doi.org/10.3390/s22208071.

Crockett, Keeley, et al. "Building Trustworthy AI Solutions: A Case for Practical Solutions for Small Businesses." IEEE Transactions on Artificial Intelligence, vol. 4, no. 4, Aug. 2023, pp. 778–91, https://doi.org/10.1109/TAI.2021.3137091.

Eboigbe, Emmanuel O., et al. "Business Intelligence Transformation Through AI and Data Analytics." Engineering Science & Technology Journal, vol. 4, no. 5, Nov. 2023, pp. 285–307, https://doi.org/10.51594/estj.v4i5.616.

Enholm, Ida Merete, et al. "Artificial Intelligence and Business Value: A Literature Review." Information Systems Frontiers, vol. 24, no. 5, Oct. 2022, pp. 1709–34, https://doi.org/10.1007/s10796-021-10186-w.

Farayola, Oluwatoyin A., et al. "Innovative Business Models Driven by AI Technologies: A Review." Computer Science & IT Research Journal, vol. 4, no. 2, Nov. 2023, pp. 85–110, https://doi.org/10.51594/csitrj.v4i2.608.

Fernandez, Joaquin. "The Leading Generative AI Companies." IoT Analytics, 4 Mar. 2025, iot-analytics.com/leading-generative-ai-companies/?utm_source=chatgpt.com.

Gomes, Poliana, et al. "Artificial Intelligence-Based Methods for Business Processes: A Systematic Literature Review." Applied Sciences, vol. 12, no. 5, Feb. 2022, p. 2314, https://doi.org/10.3390/app12052314.

Kaggwa, Simon, et al. "AI in Decision Making: Transforming Business Strategies." International Journal of Research and Scientific Innovation, vol. X, no. XII, 2024, pp. 423–44, https://doi.org/10.51244/IJRSI.2023.1012032.

Kitsios, Fotis, and Maria Kamariotou. "Artificial Intelligence and Business Strategy towards Digital Transformation: A Research Agenda." Sustainability, vol. 13, no. 4, Feb. 2021, p. 2025, https://doi.org/10.3390/su13042025.

Meske, Christian, and Enrico Bunde. "Transparency and Trust in Human-AI-Interaction: The Role of Model-Agnostic Explanations in Computer Vision-Based Decision Support." Artificial Intelligence in HCI, edited by Helmut Degen and Lauren Reinerman-Jones, Springer International Publishing, 2020, pp. 54–69, https://doi.org/10.1007/978-3-030-50334-5_4.

Schmitt, Marc. "Automated Machine Learning: AI-Driven Decision Making in Business Analytics." Intelligent Systems with Applications, vol. 18, May 2023, p. 200188, https://doi.org/10.1016/j.iswa.2023.200188.

AI and Dental Caries: of Detection, Prediction, and Clinical Implementation
By Jason Noh

Abstract

The adoption of artificial intelligence (AI) in dentistry for the identification, diagnosis, and prediction of dental caries has been growing at a rapid rate due to its capabilities to perform across multiple forms of data and modalities [1,8,25]. A total of 36 studies were synthesized using this systematic review regarding AI applications for caries-related objectives such as detection, segmentation, clinical decision support, educationally-based uses, caries-risk prediction, and health economic evaluations. Studies included in this systematic review utilized 2D radiographs, intraoral photographs, cone-beam computed tomography (CBCT), intraoral scans, and electronic health records (EHRs). Following PRISMA guidelines, we evaluated and identified studies that provided performance metrics for AI systems or provided summary evidence on such, and extracted information on the model type, training size, data source, evaluation methodology, clinical objective, comparisons with humans, stage of development, and ethics reporting.

Most of these models employed deep learning techniques, primarily convolutional neural network (CNN) based architectures [3,7,20] with ResNet variants, U-NET, YOLO, and MobileNet being the most commonly employed architectures. Nearly 76% of studies focused solely on caries detection, while other studies also examined multi-condition assessments, health economics, treatment outcomes predictions, and education-based applications. Pooled sensitivities of 0.76–0.94 and pooled specificities of 0.87–0.95 were reported from meta-analyses conducted within systematic reviews, although with considerable variability ($I^2 > 80–98\%$) [1,6,8]. Sensitivities of primary studies reporting sensitivity varied from 0.36 (dental clinicians) to 0.99 (AI-based system), while having a median value of 0.84 and a mean of 0.82. Specificity values reported in primary studies varied from 0.63 to 0.99 (with a median of 0.91 and a mean of 0.90). The Area Under the Receiver Operating Characteristic curve (AUROC) reported in nearly all of these studies was generally within the 0.73–0.98 range, indicating that many models could effectively discriminate between caries and non-caries cases [4,12,35].

These findings clearly show differences in performance dependent upon lesion severity (early noncavitated lesions: 20–72% sensitivity vs. advanced cavitated lesions: 71–99% sensitivity, representing a 1.2–4.4 fold difference) and tooth surface location (11–94% sensitivity with anterior and proximal surfaces representing 11–58% vs. posterior occlusal surfaces representing 88–94%) [6,10,17]. Although 3/20 studies (15%) had externally validated their results, with variations in performance ranging from −34% to +24% [13,15,33]. AI models have shown significant improvements over those developed by junior dentists, with a 4–108% relative sensitivity gain, but decreased specificity [31,33,35]. Additionally, two analyses demonstrated that AI models are cost-effective with a resulting average of 2.0–2.4 years of tooth retention and reductions in costs of €24–41 per patient when compared to conventional methods; particularly when less invasive treatment options are implemented [30,31].

However, there exist several limitations, including limited sample sizes from single-center studies (54% <3000 images); lack of external validation (15%); lack of reporting demographics related to age/sex/ethnicity (80% reported no demographic data); and very little utilization in real-world clinical environments despite high accuracy rates reported. Future work will need to be directed toward larger and more diverse datasets; increased testing in real-world clinical settings; external validations; standardization of evaluation methodologies; explicit fair-checking mechanisms for AI systems; incorporation of multimodal approaches; and clinically-integrated workflow design in order to make research applicable in actual clinical practice.

Introduction

Dental caries is one of the most common chronic conditions worldwide. Approximately 2.4 billion individuals have untreated caries in their permanent teeth, while approximately 621 million children experience untreated caries in their primary teeth [37]. The early detection of various lesions is necessary for non-invasive and preventative treatments such as fluoride application, dietary adjustments, and enhanced oral hygiene practices. Despite this necessity, traditional radiographic and visual examinations suffer from considerable variability among different clinicians (kappa coefficients ranging from 0.40-0.65), and have limited sensitivity for detecting early approximal lesions on bitewing radiographs (clinicians miss 30%-50% of these lesions) [6,10,26].

The recent advancements made in artificial intelligence, specifically with deep learning utilizing CNNs, have led to developing automated tools to assist with caries detection and associated diagnostics based upon radiographs, intraoral scans, photographs, and other forms of multimodal data [3,7,20]. CNNs can automatically learn hierarchical features directly from raw input data, eliminating the requirement for human "engineering" and performing better than humans on certain pattern recognition tasks [7,13,24]. Several systematic reviews and umbrella reviews previously summarized subsets of the literature; however, they focus primarily on either individual modalities (e.g., bitewing radiographs, panoramic radiographs, intraoral images, etc.) or simply the diagnostic accuracy of these systems [4,8,24]. Therefore, a comprehensive synthesis addressing various aspects, including imaging types, problem types (diagnostics/predictive models), stage of development (prototype/commercially available), issues surrounding stratifying lesion severity, surface-specific performance, barriers to translating results to practice, providing health economic evidence, and examining equity concerns, remains an unmet need within the existing literature.

Ultimately, this systematic review will provide a comprehensive overview of how AI has been utilized for caries-related tasks in dentistry including model types; training set sizes; clinical target populations; source of data; evaluation methodologies; performance measures (broken down into lesion severity and tooth surface); clinician comparisons; level of development; transparency concerning privacy/consent/fairness of each of the 36 included studies.

Methods

1. **Research question overview (3-4 smaller questions)**

How can AI effectively be used to detect and diagnose dental caries, and what are the current limitations regarding diagnostic accuracy, real-life clinical use, and ethical concerns such as data privacy and biases?

How can AI be used to improve the treatment and diagnosis of dental caries?

2. **Pubmed search strategy**

We will use PubMed to research our question.

3. **Screening protocol**

We will perform a two-stage screening process, first screening the title and abstract, followed by the full text of the manuscripts obtained through our search method.

4. **Inclusion criteria**
    - Population: We are not limiting our research to a specific population
    - Concept: The key concepts of our paper are Dental caries and AI, screening for caries, technological restrictions using AI, and possible hazards using AI.
    - Context: We are not limiting our context to a specific demographic context
    - Source of Evidence: PubMed
    -

5. **Flow Chart/Data Extraction**

    The PRISMA flowchart will be utilized to visually represent the numerical results of the reviews and the inclusion-decision process. Our flowchart will clearly outline the steps involved in finding studies, eliminating duplicates, selecting the correct research, retrieving full articles from the library, and presenting the final analysis. Additionally, we will create a charting template to observe the data extraction.

Results

Geographic Distribution

The researchers of the 36 included studies came from over 20 different countries from six continents. The majority of studies were from China, with a total of 6 (17%). Next in line was Germany with a total of 5 (8%); then Denmark with a total of 3 (8%), then India with a total of 3 (8%), and the USA with a total of 4 (11%) studies. Each of the following countries also had 1–2 studies as part of this literature review: South Korea, Hungary, the Netherlands, Spain, Italy, the UAE, Egypt, Iran, Qatar, the UK, Sweden, Indonesia, and Cambodia.

Aim Overview

Of all 36 of the studies, 75% (27/36) were primarily concerned with the detection and diagnosis of caries via the use of intra-oral images (radiographs, panoramic images), intra-oral scans, or CBCT. Most studies have evaluated deep learning systems for detection, classification, and segmentation of dental caries [4,11,19]. Systematic reviews made up 45% (16/36) of the reviewed studies [1,8,25].

The remaining studies had other primary objectives but were still concerned with caries. For example, many evaluated multiple conditions simultaneously, such as caries + calculus + gingivitis; or caries + sealants [21,35]; some examined prediction of treatment outcome, such as pulp exposure risk [28], and a few examined cost effectiveness [30,31], educational impact on students via AI [29], and usability of smartphone application [2].

As shown in 27 studies (75%), the most prominent problem was the detection of dental caries, including approximal caries [6,10,17]; occlusal caries [11,16,19]; early childhood caries (ECC) [2,15,32]; third molar caries [34]; secondary caries [5]. More specific applications included lesion severity classification (six studies, 17%) [6,11,15], segmentation (twelve studies, 33%) [11,13,21], and risk prediction (three studies, 8%) [18,20,23]. Two studies (6%) conducted economic evaluations [30,31], one study (3%) evaluated educational effectiveness [29], and one study (3%) assessed an app that directly confronts patients [2].

Type of AI Used

Most of the 36 studies used deep learning, especially CNNs [3,7,25]. ResNet variants were the most used architecture (12/36, 33%) [11,16,33]. U-Net showed up in 10 studies (28%), especially in segmentation tasks, including standard U-Net [13,19,21], Attention U-Net [15], and a new Oral-Mamba variant, which combined U-Net with State Space Models [21]. YOLO (You Only Look Once) models were used in 8 studies (22%), mainly for object detection in radiographs and photographs, including YOLOv3, YOLOv5, YOLOv7, YOLOv8, YOLOX, and YOLOv7 + CBAM with attention mechanisms [5,11,12].

MobileNet variants (MobileNetV2, MobileNetV3) appeared in 7 studies (19%), predominantly for smartphone-based applications and edge deployments that require lightweight models [2,9,17]. Inception variants (Inception-v3, Inception-ResNet-v2) were used in 5 studies (14%), often as parts of ensemble systems [11,16,33]. For instance, segmentation and multi-stage detection tasks, Mask R-CNN, Faster R-CNN, and Cascade R-CNN appeared in 6 studies (17%) [11,16,33]. DenseNet and EfficientNet families each appeared in 3 studies (8%) [9,11,27]. VGG architecture (VGG-16, VGG-19) was used in 4 studies (11%) [11,19,34].

Novel architecture included Oral-Mamba (combining State Space Models with U-Net for 25% faster inference) [21], ToothNet (custom mutli-task architecture with parallel classification heads) [35], and mentions of Transformer-based models in reviews, although they were not yet used in primary studies [20,23,25] Traditional machine learning methods (SVM, Random Forest, Decision Trees) appeared in 5 studies (14%), primarily in reviews summarizing older literature or as baseline comparisons [18,22,26].

Ensemble learning approaches were used in 2 studies (6%), with Kang et al. 2024 demonstrating that combining ResNet-50, Inception-v3, Inception-ResNet-v2, and Faster R-CCN improved AUC from 0.91–0.94 (individual models) to 0.925 (ensemble) [16]. Transfer learning using ImageNet or COCO pre-training was employed in 18/20 primary studies (90%) [9,11,16], with only two studies training from scratch [30,33].

Training Size

Training set sizes significantly varied across primary studies. Approximately 54% of primary studies used fewer than 3,000 images, while 15% used more than 10,000 images. The smallest dataset consisted of 87 images (reported in Prados-Privado 2020 review of early neural network studies) [26], while the largest was over 150,000 images [17,36].

Four studies (20%) used fewer than 500 images (100 radiographs [12], 229 radiographs [27], 292 radiographs [28], and 50 [29]. Two studies (10%) used 500–1,000 images, including 332 carious teeth [15] and 500 third molar images [34]. Seven studies (35%) used 1,000–3,000 images, including 2,713 CBCT slices [11], 2,682 intraoral images [16], 2,417 intraoral photographs [19], 2,019 training images [21], and 1,020 intraoral photos [35]. Two studies (10%) used 3,000–10,000 images, including 3,686 radiographs [31] and 4,500 radiographs [33]. Four studies (20%) used or evaluated models trained on more than 10,000 images: 25,500 augmented images [9], 48,771 tooth crops [30], Karakuş 2024 (evaluated Aiyakankan A9X trained on more than 150,000 images, own dataset size not reported) [17], and Zhang 2024 (evaluated same system on 4,361 teeth, model trained on more than 150,000 images) [36]. Three studies conducted external validation only or did not report training data size [2,13,32]. Data augmentation (such as rotation, flipping, brightness adjustments) was widely applied in 17/20 primary studies (85%), though reporting was heterogeneous [9,11,13]. Augmentation typically multiplied datasets by 5–50 times, meaning reported dataset sizes may overestimate the number of unique images.

Study Orientation: Predictive vs. Diagnostic

Of the 36 studies, 30 (83%) were classified as diagnostic only, focusing on the detection or classification of existing carious lesions [4,11,19]. Six studies (17%) had predictive components: Khanagar 2022 review identified 8 primary studies on caries risk prediction across 23 years of literature [18]; Liang 2025 and Marwaha 2025 reviews discussed risk prediction conceptually [20,23]; Ramezanzade 2023 predicted pulp exposure risk before caries excavation (the only primary predictive study) [28]; and Schwendicke 2021 and Schewendicke 2022 used Markov models to project long-term tooth retention and costs [30,31].

This distribution of data demonstrates that the vast majority of work focuses on current disease detection and classification rather than future risk prediction, despite the importance of predicting which patients will develop caries. The limited predictive evidence base (1 primary study and 3 reviews mentioning prediction) represents a major gap given the preventive care paradigm shift in modern dentistry [18,20,28].

Types of Problems Addressed

The most commonly studied problem was dental caries detection, present in 27 of 36 studies (75%), including:

- Approximal (proximal) caries on posterior teeth (12 studies, 33%) [6,10,17]
- Occlusal caries (11 studies, 31%) [5,11,19]
- Early Childhood Caries (ECC) in primary teeth (3 studies, 8%) [2,15,32]
- Third molar (wisdom tooth) caries (1 study, 3%) [34]
- Secondary caries beneath fixed dental prostheses (1 study, 3%) [5]
- Smooth surface caries (labial, buccal, lingual, palatal) [17,19,21]

Lesion severity classification appeared in 6 studies (17%), utilizing ICDAS (International Caries Detection and Assessment System), ICCMS (International Caries Classification and Management System), or custom systems to categorize lesions as enamel-only, dentin-involved, cavitated, or by depth/severity [6,11,15].

Segmentation tasks (pixel-level or lesion-level delineation) showed up in 12 studies (33%) [11,13,21]. Multi-condition assessment was novel in 2 studies (6%): Liu 2024 simultaneously detected caries, dental calculus, and gingivitis (IoU 0.64-0.71) [21], and Xiong 2024 detected caries and fissure sealants (AUC 0.925 and 0.902, respectively) [35]. Treatment outcome prediction appeared in a study (3%): Ramezanzade 2023 predicted pulp exposure risk before excavation with AUC 0.73, sensitivity 0.62, specificity 0.83 [28]. Health economic evaluation was conducted in 2 studies (6%): Schwendicke 2021 and 2022 both demonstrated AI cost-effectiveness with a dominant strategy (better outcomes, lower costs) under minimally invasive paradigms [30,31].

Educational effectiveness was evaluated in 1 study (3%): Schropp 2024 found that providing AI assistance (AssistDent) to dental students during training did not improve diagnostic accuracy and showed no learning benefit [29]. Usability assessment appeared in 1 study (3%): Al-Jallad 2022 evaluated the AICaries smartphone app for at-home ECC screening by parents, achieving a 78.4 System Usability Scale score (excellent) with 78.6% producing diagnostic-quality photos [2].

Data Sources and Modalities

There was great variety in the use of various image acquisition methods that have been studied. There were 23 studies (64%) that used 2D radiographic images alone; there were 12 studies (33%) that used bitewing radiographs alone [6,10,17], 8 studies (22%) that used panoramic radiographs alone [5,14,34], 6 studies (17%) that used periapical radiographs alone [28,30,31] and 5 studies (14%) that used combinations of these three types of radiographic images [5,28,30]. Only three studies (8%) used 3D CBCT; however, only one study (Chen et al., 2024) analyzed lesion severity based on axial slices using CBCT [11].

11 studies (31%) included intraoral photographs taken either by professional cameras [2,13,19], intraoral cameras [2,13,19], or smartphone cameras. The photographs taken were

sourced via internet for Frenkel et al., 2024 (718 photographs) [13]. Two studies (6%) included intraoral 3D scans [15].

4 studies (11%) included smartphone photography. These studies were Al-Jallad et al., 2022 (photographed by parents of children's teeth) [2]; Boy et al., 2025 (optimized for mobile deployment) [9]; Dhanak et al., 2024 (used real time video detection at 14 frames per second) [12]; and Xiong et al., 2024 (both camera and telephone photographs) [35].

Two studies (6%) referenced EHR data or clinical metadata. Liang et al., 2025 referenced multimodal integration of imaging data, patient demographic data, genetic information, and microbiome when developing their model [20]. Vinayahalingam et al., 2021 used electronic medical records (EMRs) as source of "ground truth" labels [34]. Various advanced imaging methods such as optical coherence tomography (OCT), near infrared light transillumination (NILT), quantitative light induced fluorescence (QLF), photothermal imaging, etc., have been reviewed, but none were the primary imaging method in any of the primary studies [18,20,23].

Only 8 studies (22%) used multiple-modal approaches. Most of these studies compared the different modalities against each other. Ramezanzadeh et al. 2023 combined the use of bitewing radiographs with other variables (patient age, patient-reported pain, treatment type) using a multi-path neural network architecture [28].

Summary of Performance Metrics

Meta-analyses from systematic reviews

Four systematic reviews conducted meta-analyses, providing estimates with 95% confidence intervals:

- Abbott et al. 2025 (7 studies): Pooled sensitivity 0.76 (95% CI 0.65–0.85), specificity 0.91 (0.86–0.95), AUC 0.92 (0.89–0.94), $I^2 > 98\%$ (very high heterogeneity) [1]
- Ammar & Kühnisch 2024 (bitewing-specific): Overall caries sensitivity 0.87 (0.76–0.94), specificity 0.89 (0.75–0.96); dentin caries 0.83/0.84; enamel caries 0.72/0.73 [6]
- Carvalho et al. 2024 (approximal caries, 21 studies): Pooled sensitivity 0.94 (0.78–0.99), specificity 0.91 (0.84–0.95) [10]
- Arzani et al. 2025 (umbrella review, 20 of 137 studies in meta): Sensitivity 0.85 (0.83–0.93), specificity 0.90 (0.85–0.95), AUC 0.86 [8]

All meta-analyses reported extreme heterogeneity ($I^2 > 80\%$), with Abbott and Arzani exceeding 98%, meaning the included studies varied substantially across studies in datasets, architectures, evaluation protocols, and caries definitions, limiting the reliability of pooled estimates [1,8,10].

Primary studies

Out of 36 total studies, 18 primary research studies reported sensitivity values:

- Range: 0.36 (human dentists in Schwendicke 2021 [31]) to 0.99 (AI models [19])
- AI sensitivity range: 0.57-0.99

- Median: 0.84
- Mean: 0.82
- Studies exceeding 0.90 sensitivity: 8/18 (44%) [9,11,16]

Specificity was reported in 18 primary studies:

- Range: 0.63–0.99
- Median: 0.91
- Mean: 0.90
- Studies exceeding 0.90 specificity: 12/18 (67%) [11,15,17]

AUROC was reported in 14 primary studies:

- Range: 0.73–0.98
- Median: 0.92
- Mean: 0.90
- Studies exceeding 0.90 AUC: 10/14 (71%) [15,16,19]

Performance by lesion severity (critical finding)

Six studies found dramatic performance differences between early and advanced caries:

- Jones et al. 2025 (intraoral scans, pediatric): Initial/moderate lesions 0.20-0.44 sensitivity vs. extensive lesions 0.71–0.88 (1.8–4.4 fold difference) [15]
- Frenkel et al. 2024 (photographs): Greyish/microcavity 0.43 sensitivity vs. noncavitated 0.93 (2.2 fold difference) [13]
- Schwarzmaier et al. 2024 (photographs, ECC): Greyish/micro 0.69 vs. destructed 0.93–0.99 (1.3-1.4 fold difference) [32]
- Ammar & Kühnisch 2024 (bitewing meta-analysis): Enamel 0.72 vs. dentin 0.83 (1.2 fold difference) [6]
- Carvalho et al. 2024 (bitewing meta-analysis): Enamel 0.72 (meta) [10]
- Chen et al. 2024 (CBCT): 0.57 sensitivity for mixed severity lesions vs. an interpretable model with segmentation had 0.92 sensitivity [11]

Pattern: Advanced cavitated lesions consistently showed a 1.2–4.4 fold higher sensitivity than early noncavitated lesions across all imaging modalities [6,11,13].

Performance by tooth surface

Karakuş et al. 2024 evaluated the Aiyakankan A9X system (MobileNet-V3 + U-Net) on bite-wing radiographs across 11 tooth surfaces (the dataset size was not reported). The study looked into interproximal, occlusal, and secondary caries detection using a deep learning approach in a single-center Chinese cohort [17]:

- Highest sensitivity: Molar occlusal 0.944, Premolar occlusal 0.941

- Moderate sensitivity: Molar palatal 0.895, Molar buccal 0.881, Molar proximal 0.586
- Low sensitivity: Incisal 0.500, Anterior palatal/lingual 0.500, Premolar proximal 0.439
- Extremely low sensitivity: Labial 0.333, Anterior proximal 0.111, Premolar buccal 0.111
- Overall clinical sensitivity: 0.8131 (average ≈ 0.589)
- Overall specificity: 0.9565 (per-surface range 0.8156–0.9968, average ≈ 0.947)
- Overall accuracy: 0.9340
- NPV: 0.9649

Evaluation Protocols

Train/test splits

Out of the 20 primary research studies, train/test split ratios were:

- 80/20 split: 6 studies (30%) [16,19,21]
- 90/10 split: 3 studies (15%) [11,27,30]
- 70/30 split: 2 studies (10%) [12,31]
- Three-way splits (train/val/test): 8 studies (40%) with variable ratios [2,5,13,15,28,29,32,33,34]
- Not clearly reported: 1 study (5%) [17]

Cross-validation

- 5-fold CV: 4 studies (20%) [5,12,19]
- 10-fold CV (stratified): 2 studies (10%): Ramezanzade 2023 [28], Schwendicke 2022 [30]
- k-fold CV (folds not specified): 3 studies (15%) [9,11,27]
- No cross-validation: 11 studies (55%) [2,13,15]

External validation

Only 3/20 primary studies (15%) conducted external validation:

- Frenkel et al. 2024: Tested an already existing AI model on 718 photos from the internet. The overall performance stayed high (92% sensitivity), but the performance on cavitated lesions dropped (from 95.7% to 71.2%) [13]
- Jones et al. 2025: External validation on 119 teeth from a separate dataset showed that the sensitivity for extensive lesions improved (from 71% to 88%), but declined for early lesions (from 44% to 29%) [15]
- Schwarzmaier et al. 2024: Validated the same Kühnisch model on 143 ECC photos. It showed excellent performance (97.2% sensitivity/specificity), higher than the general population [32]

All other 17 primary studies (85%) used internal validation only (single-institution train/test split), raising concerns about generalizability [5,9,11].

Human Comparisons

13 studies (36%) directly compared AI performance with human clinicians [16,27,28]:

AI superiority shown

- Schwendicke 2021: AI sensitivity 0.75 vs. dentists 0.36 (from 0.19 to 0.65); +108% relative improvement [31]
- Szabó et al. 2024: AI sensitivity 0.73–0.76 vs. humans alone 0.51; +43% relative improvement across 4 evaluation phases [33]
- Xiong et al. 2024: AI AUC 0.925 (caries) vs. 1-year dentists 0.749; +23% massive AUC improvement; F1 0.810 vs 0.599 (+35% relative) [35]
- Ramezanzade 2023: AI sensitivity 0.62 vs. dental students 0.53–0.60; +4–17% relative improvement [28]
- Vinayahalingam 2021: AI AUC 0.90 vs. cite prior clinician study 0.768 [34]

- AI specificity is around 0.83–0.88 [28,30,31]
- Human specificity is around 0.91–0.96 [29,31,33]
- Pattern: AI had higher sensitivity, humans had higher specificity [28,29,31]

Development Stage

Of the 36 studies

- Research prototypes: 16 studies (44%), early validation phase [1,3,4]
- Clinical evaluation: 3 studies (8%) with prospective real-world testing (Zhang 2024, Schropp 2024, Ramezanzade 2023) [28,29,36]
- Commercial products: 4 studies (11%) evaluating deployed systems:
    - Aiyakankan A9X (MobileNet-v3 + U-Net): Karakuş 2024 (Turkish researchers, Chinese dataset, dataset size not reported) [17], Zhang 2024 (Chinese researchers, Chinese dataset with 4,361 teeth from 191 patients) [36]
    - Diagnocat (Mask R-CNN + Cascade R-CNN): Szabó 2024 (Hungary/Turkey) [33]
    - AssistDent: Schropp 2024 (Denmark/UK, educational failure) [29]
    - AICaries: Al-Jallad 2022 (USA, usability only, no performance metrics) [2]
- Publicly accessible models: 3 studies (8%):
    - Kühnisch/Felsch dental-ai.de Model: Kühnisch 2022 [19], Frenkel 2024 [13], Schwarzmaier 2024 [32]
    - Xiong ToothNet [35]
- Development stage not specified: 10 studies (28%) [14,18,20]

This distribution implies that most of the models are still in early or intermediate stages of validation, with widely deployed tools being rare. The three studies using the publicly

accessible dental-ai.de model show an uncommon level of reproducibility in dental AI research [13,19,32].

Privacy and Consent Reporting

Ethics approval: Reported in 16/20 primary studies (80%), with IRB (Institutional Review Board) identifiers shown in 12 studies (60%) [2,11,13]. Four studies (20%) did not report ethics approval [9,16,34]. Karakuş 2024 reported ethics approval with written informed consent included [17].

Informed consent: Obtained in 14 studies (70%) [2,11,13]. However, waived for anonymized retrospective data in 6 studies (30%) [9,16,30].

Data anonymization/de-identification: Explicitly reported in 19/20 primary studies (95%), with methods such as removal of patient identifiers, encryption, and secure storage [5,9,11].

Demographic reports

- Age reported: 4 studies (20%) [2,15,28]
- Sex/gender reported: 3 studies (15%) [15,17,28]
- Race/ethnicity reported: 0 studies (0%)
- Socioeconomic status reported: 1 study (5%): Al-Jallad 2022 targeting low-income families [2]
- None reported: 12 studies (60%) [5,9,11]

Algorithmic fairness analysis

- Explicit fairness assessment: 0 studies (0%)
- Performance stratified by demographic subgroups: 0 studies (0%)
- Bias discussed qualitatively in limitations: 20 studies (100%) [2,5,9]
- Bias mitigation strategies implemented: 1 study[27]

Privacy-protecting techniques

- Federated learning: 0 studies implemented; 1 recommended (Marwaha 2025) [23]
- Differential privacy: 0 studies
- Secure multi-party computation: 0 studies

Despite 95% reporting basic privacy measures (anonymization, IRB approval), no studies assessed algorithmic fairness or performance across demographic subgroups, although there were known health disparities in caries prevalence by race, ethnicity, and socioeconomic status.

Summary, Limitations, and Future Directions

Summary

This systematic review demonstrates that AI systems for caries detection often demonstrate high diagnostic performance, especially the CNN-based models trained on radiographic and photographic data, with pooled sensitivities of 0.76–0.94, specificities of 0.87–0.95, and AUROC often between 0.84–0.98 [1,8,10]. However, performance varies significantly by lesion severity (early lesions with 20–72% sensitivity compared to advanced lesions with 71–99%, showing a 1.2–4.4 fold difference). Tooth surface location also had an impact on performance (11–94% sensitivity range) [6,11,17]. AI consistently outperformed junior dentists by a 4–108% relative sensitivity improvement but showed lower specificity trade-offs (83–88% vs. 91–96%) [28,31,33].

Two health economic analyses demonstrated AI cost-effectiveness as a better strategy (better outcomes and lower costs), using minimally invasive treatment paradigms, with 2.0 -- 2.4 years of tooth retention gains and €24-41 savings per patient [30,31]. However, this depends on treating early AI-detected cavities with fluoride, not filling. Hence, if all detected cavities are filled, AI becomes more costly and less effective due to unnecessary treatment [30,31].

Limitations

Common biases and limitations are shown across the 36 included studies. A major limitation is that many models were trained on small, single-center datasets, with around 54% using fewer than 3,000 images and only 15% exceeding 10,000 images. This raises concerns about overfitting and limiting generalizability. Another reason for uncertainty is the limited external validation, since only 15% of primary studies were validated on independent datasets, with performance changes varying from -34% to +24% across populations [13,15,33]. Furthermore, demographic diversity was almost fully absent, with 80% of the studies reporting no age, sex, or ethnicity data, and no studies assessing algorithmic fairness across subgroups despite known health disparities.

Additional weaknesses include heterogeneous evaluation protocols with inconsistent train/test splits, 55% of students lacking cross-validation, and 95% lacking temporal validation. The quality of ground truth labels was also problematic, with 72% not reporting whether different dentists agreed on the diagnoses, and 95% relying on visual examination rather than microscopic examination. Some studies did not include cases where dentists disagreed, making results appear better by removing difficult cases. All meta-analyses reported extreme heterogeneity (I^2 > 80-98%), and publication bias was detected [8].
Critical performance gaps appeared, including an 8.5-fold difference (94% vs 11% sensitivity), revealing limitations for surfaces that have poor access [17,36], and early lesion detection failure with 20–72% sensitivity for noncavitated lesions, which is insufficient for preventive care applications [6,10,13]. Despite high reported accuracies commonly exceeding 90%, most systems are prototypes with limited clinical deployment.

Future directions

Future research should focus on several important areas. Larger multi-institutional datasets consisting of diverse demographics, imaging equipment, and clinical settings are needed to address concerns related to generalizability. External validation through multi-center prospective studies across different populations, equipment, and protocols needs to become more common [13,15,33]. Future studies need standardized evaluation methods such as consistent caries classification systems (like ICDAS), clear standards for labeling, and specific definitions of what is being measured.

Explicit fairness assessments should be mandatory, with performance graded by age, sex, race, ethnicity, socioeconomic status, insurance status, and geographic region. Future studies should also utilize algorithmic fairness constraints such as demographic parity and equalized odds during development. Multi-modal integration combining radiographs, photographs, clinical data, and EHR should be looked into to overcome single-modality limitations, alongside advanced techniques such as QLF, NILT, OCT, and fluorescence imaging [18,20,23]. Methods specifically optimized for early noncavitated lesion detection should be developed to enable preventive care models [6,10,13].

Studies related to real-world implementation are needed, including clinical trials, workflow integration assessments, cost-effectiveness evaluations in healthcare systems beyond Germany, and user-centered design research. Advanced methodologies should be explored, including semi-supervised learning to reduce annotation burden, federated learning for privacy-preserving multi-institutional training, multi-task learning for comprehensive oral health assessment, and ensemble methods for improved robustness [16,21,23]. Health economic evaluations should extend beyond Germany to examine the cost-effectiveness of photo-based, CBCT-based, and smartphone-based approaches, with analyses of AI access and impact on health disparities [30,31].

Finally, addressing limitations specific to surfaces requires recognizing that multi-modal approaches are necessary for comprehensive assessment, with radiographs being essential for detecting approximal caries, regardless of photo-based AI results [17,36]. Suggestions for commercial systems like Aiyakankan A9X include integrating multimodal imaging such as fluorescence and advanced radiographic techniques, improving detection of early lesions, increasing dataset diversity across populations and imaging conditions, combining patient-generated images for telemedicine applications, and conducting long-term clinical usefulness evaluations in diverse practice settings [17,36].

Discussion

Summary of Evidence

This study aims to systematically map the application of AI in dental caries and related conditions across 36 studies recorded in multiple countries, imaging modalities, and clinical tasks. The most important finding was that AI systems, particularly CNN-based models for the detection of caries, consistently achieved high diagnostic performance, with median sensitivities

and specificities around or above 0.85 and many reaching 0.90–0.98, including AUROC values up to 0.98 [1,4,8].

However, the most significant discovery is the dramatic variation in performance by lesion severity and tooth surface location. Early noncavitated lesions showed a 20–72% sensitivity, compared to 71–99% for advanced cavitated lesions, showing a 1.2–4.4 fold performance gap [6,10,13].

Another important finding is that only 15% of primary studies conducted external validation, with performance changes ranging from -34% to +24% across different populations [13,15,33]. One surprising result was that providing AI to dental students during training did not improve diagnostic accuracy [29], suggesting potential risks related to AI dependency rather than learning enhancement.

Comparison With Previous Work

These results support the findings of many previous works in this area. This result is consistent with several recent systematic reviews and meta-analyses, which similarly report high sensitivity (0.76–0.94) and specificity (0.87–0.95) for AI models in caries detection across bitewing, panoramic, and periapical radiographs [5,7,10].

However, the findings of this study build on prior work by quantifying the 1.2–4.4 fold performance gap between early and advanced lesions, as well as the 8.5-fold variation across tooth surfaces (11–94% sensitivity range), which were not systematically examined in previous reviews [6,11,17]. This study has been unable to show that AI performs adequately for early preventive screening, in contrast to positive conclusions from earlier reviews that did not account for differences in lesion severity and variation across anatomical location [5,7,25].

Limitations of the Evidence Base

Despite generally strong performance metrics, several limitations are present across the 36 included studies. A significant limitation is that many models were trained on small, single-center datasets, with about 54% using fewer than 3,000 images and only 15% of studies exceeding 10,000 images [11,16,19]. These data could infer that performance may not be reproducible across different populations, imaging protocols, and devices [4,8,24].
Another reason for uncertainty is the limited external validation. It is important to acknowledge the possible biases in these results, as only 15% of primary studies evaluated models on independent datasets [13,15,33]. Therefore, the results need to be interpreted with caution, given the -34% to +24% performance variability observed across validation sites [13,15,33].

Furthermore, although privacy measures and IRB approvals were widely reported (80% of primary studies), explicit assessment of demographic bias and fairness was almost entirely absent. It is possible that these results may not be reproducible on a wide scale across diverse demographic groups, despite known health disparities in caries prevalence by race, ethnicity, and socioeconomic status.

Clinical Implications

The findings suggest that AI currently works well for screening cavitated posterior lesions on bitewing radiographs and intraoral photographs (88–94% sensitivity), but fails for early preventive screening (20–72% sensitivity) [6,13,15]. An implication of this finding is the possibility that AI should be deployed as an assistive screening tool rather than for diagnostic confirmation, given the moderate positive predictive value (77–88%) requiring clinical verification [1,10,17].

One of the issues that emerges from these findings is the critical dependency on treatment philosophy. The results of the health economic analyses suggest that AI cost-effectiveness (€24–41 savings, 2.0 -- 2.4 years tooth retention gains) depends entirely on non-restorative management of early lesions [30,31]. If clinicians restore all AI-detected lesions, AI becomes more costly and less effective through overtreatment [30,31].

Methodological Innovations

One interesting finding is that semi-supervised learning matched competitive benchmarks (mIoU 56.5%) with needing 94% less labeled data than typical studies, demonstrating a way to reduce annotation costs [27]. Another important finding is that multi-task learning allowed for comprehensive oral health assessment, with systems simultaneously detecting caries, calculus, and gingivitis while running 25% faster than baseline models [21,35].

These results provide further support for the hypothesis that ensemble methods systematically improve robustness, with combined models increasing AUC from 0.91–0.94 (individuals) to 0.925 (ensemble) [16]. However, with computational costs increasing 4-fold, caution must be applied, as this approach might not be suitable for real-life clinical applications [16].

Equity and Fairness Considerations

Very little was found in the literature on algorithmic fairness in dental caries AI systems. In reviewing the 36 included studies, no data were found on performance stratification by race, ethnicity, or socioeconomic status, despite known 2–3 fold higher untreated caries rates among Black and Hispanic children compared to white children in the United States.

It seems possible that these results are due to sampling bias, with single-center tertiary care institutions underrepresenting community clinics and underserved populations. The possible interference of equipment bias, phenotypic variation, and labeling bias cannot be ruled out. These findings suggest that explicit fairness assessments should be mandatory before clinical deployment to prevent exacerbating existing health inequities.

Future Directions

Future work is required to address the following gaps. Additional research is needed to better understand early lesion detection mechanisms, as systems currently only show a 20–72% sensitivity, which is insufficient for preventive care paradigms [6,13,15]. Further studies should

prioritize multi-institutional external validation to examine how AI systems perform across different patient demographics and clinical settings. [13,15,33].

A future study with more focus on multi-modal integration is suggested, combining radiographs, photographs, clinical data, and EHR to overcome the limitations of single-modality [20,23,28]. Further work is needed to develop methods specifically optimized for poorly accessible tooth surfaces, which show a sensitivity of around 11–58% on photographs [17,36]. Suggested improvements for commercial systems include the addition of multimodal imaging (fluorescence, advanced radiographic techniques), improving the detection of early and subtle lesions, increasing dataset diversity across populations and imaging conditions, and evaluating long-term clinical usefulness in diverse settings [17,36].

Despite promising results for advanced lesion detection, there are still some concerns about real-world application, workflow integration, and its long-term clinical impact. To fully understand the role of AI in managing caries, additional studies will be necessary that directly address that preserves patient privacy, standardized evaluation protocols, and equity analyses.

Conclusion

This systematic review synthesized 36 studies on AI applications in dental caries detection, prediction, and related tasks across radiographic, photographic, intraoral scan, smartphone, and multimodal data sources. Overall, deep learning models, in particular CNN-based models, showed a high diagnostic performance, with many achieving high sensitivities and specificities above 0.90 along with AUROC values between 0.84 and 0.98 for tasks related to caries [1,4,8].

However, the current evidence is limited by small, single-center datasets, limited external validation, inconsistent evaluation protocols, and a lack of reporting on fairness and demographic bias [5,7,10]. Hence, future work should focus on larger, more diverse datasets, multi-center validation, and clinically integrated implementations that address privacy, ethics, and equity [18,21,22].

Taken together, this review suggests that AI has high potential to enhance dental caries detection and management, yet it will require careful methodological care, transparent reporting, and close collaboration between clinicians, researchers, and technologists to ensure safe, effective, and fair deployment in everyday dental practice.

Works Cited

[1] Abbott, L. P., Saikia, A., & Anthonappa, R. P. (2025). Artificial intelligence platforms in dental caries detection: A systematic review and meta-analysis. Journal of Evidence-Based Dental Practice, 25(1), 102077. https://doi.org/10.1016/j.jebdp.2024.102077

[2] Al-Jallad, N., Ly-Mapes, O., Hao, P., Ruan, J., Ramesh, A., Luo, J., Wu, T. T., Dye, T., Rashwan, N., Ren, J., Jang, H., Mendez, L., Alomeir, N., Bullock, S., Fiscella, K., & Xiao, J. (2022). Artificial intelligence-powered smartphone application, AICaries, improves at-home dental caries screening in children: Moderated and unmoderated usability test. PLOS Digital Health, 1(6), e0000046. https://doi.org/10.1371/journal.pdig.0000046

[3] Al-Khalifa, K. S., Alsubaie, F. A., & Almashhadani, S. S. (2024). The use of artificial intelligence in caries detection: A review. Bioengineering, 11(9), 936. https://doi.org/10.3390/bioengineering11090936

[4] Albano, D., Galiano, V., Basile, M., Di Luca, F., Gitto, S., Messina, C., Cagetti, M. G., Del Fabbro, M., Tartaglia, G. M., & Sconfienza, L. M. (2024). Artificial intelligence for radiographic imaging detection of caries lesions: A systematic review. BMC Oral Health, 24, 274. https://doi.org/10.1186/s12903-024-04046-7

[5] Ayhan, B., Ayan, E., & Atsü, S. (2025). Detection of dental caries under fixed dental prostheses by analyzing digital panoramic radiographs with artificial intelligence algorithms based on deep learning methods. BMC Oral Health, 25, 216. https://doi.org/10.1186/s12903-025-05577-3

[6] Ammar, N., & Kühnisch, J. (2024). Diagnostic performance of artificial intelligence-aided caries detection on bitewing radiographs: A systematic review and meta-analysis. Japanese Dental Science Review, 60, 128–136. https://doi.org/10.1016/j.jdsr.2024.02.001

[7] Anil, S., Porwal, P., & Porwal, A. (2023). Transforming dental caries diagnosis through artificial intelligence-based techniques. Cureus, 15(7), e41694. https://doi.org/10.7759/cureus.41694

[8] Arzani, S., Karimi, A., Iranmanesh, P., Yazdi, M., Sabeti, M. A., Nekoofar, M. H., Kolahi, J., Bang, H., & Dummer, P. M. H. (2025). Examining the diagnostic accuracy of artificial intelligence for detecting dental caries across a range of imaging modalities: An umbrella review with meta-analysis. PLOS ONE, 20(8), e0329986. https://doi.org/10.1371/journal.pone.0329986

[9] Boy, A. F., Akhyar, A., Arif, T. Y., & Syahrial, S. (2025). Artificial intelligence for dental caries detection: A mixup, fine-tuning, and quantization approach on the MobileNetV2 model. Journal of Conservative Dentistry and Endodontics, 28(8), 764–771. https://doi.org/10.4103/JCDE.JCDE_362_25

[10] Carvalho, B. K. G., Nolden, E.-L., Wenning, A. S., Kiss-Dala, S., Agócs, G., Roth, I., Kerémi, B., Géczi, Z., Hegyi, P., & Kivovics, M. (2024). Diagnostic accuracy of artificial intelligence for approximal caries on bitewing radiographs: A systematic review and

meta-analysis. Journal of Dentistry, 151, 105388. https://doi.org/10.1016/j.jdent.2024.105388

[11] Chen, S., Yang, Y., Wu, W., Wei, R., Wang, Z., Tay, F. R., Hu, J., & Ma, J. (2024). Classification of caries based on CBCT: A deep learning network interpretability study. Journal of Imaging Informatics in Medicine, 37, 3160–3173. https://doi.org/10.1007/s10278-024-01143-5

[12] Dhanak, N., Chougule, V. T., Nalluri, K., Kakkad, A., Dhimole, A., & Parihar, A. S. (2024). Artificial intelligence enabled smart phone app for real-time caries detection on bitewing radiographs. Bioinformation, 20(3), 243–247. https://doi.org/10.6026/973206300200243

[13] Frenkel, E., Neumayr, J., Schwarzmaier, J., Kessler, A., Ammar, N., Schwendicke, F., Kühnisch, J., & Dujic, H. (2024). Caries detection and classification in photographs using an artificial intelligence-based model—An external validation study. Diagnostics, 14(20), 2281. https://doi.org/10.3390/diagnostics14202281

[14] Hung, M., Yevseyevich, D., Khazana, M., Schwartz, C., & Lipsky, M. S. (2025). Charting new territory: AI applications in dental caries detection from panoramic imaging. Dentistry Journal, 13(8), 366. https://doi.org/10.3390/dj13080366

[15] Jones, B., Lambach, M., Chen, T., Michou, S., Kilpatrick, N., Curtis, N., Burgner, D. P., Vannahme, C., & Silva, M. (2025). Dental caries detection in children using intraoral scans and deep learning: A pilot study. Journal of Dentistry, 160, 105906. https://doi.org/10.1016/j.jdent.2025.105906

[16] Kang, S., Shon, B., Park, E. Y., Jeong, S., & Kim, E.-K. (2024). Diagnostic accuracy of dental caries detection using ensemble techniques in deep learning with intraoral camera images. PLOS ONE, 19(9), e0310004. https://doi.org/10.1371/journal.pone.0310004

[17] Karakuş, R., Öziç, M. Ü., & Tassoker, M. (2024). AI-assisted detection of interproximal, occlusal, and secondary caries on bite-wing radiographs: A single-shot deep learning approach. Journal of Imaging Informatics in Medicine, 37, 3146–3159. https://doi.org/10.1007/s10278-024-01113-x

[18] Khanagar, S. B., Alfouzan, K., Awawdeh, M., Alkadi, L., Albalawi, F., & Alfadley, A. (2022). Application and performance of artificial intelligence technology in detection, diagnosis and prediction of dental caries (DC)—A systematic review. Diagnostics, 12(5), 1083. https://doi.org/10.3390/diagnostics12051083

[19] Kühnisch, J., Meyer, O., Hesenius, M., Hickel, R., & Gruhn, V. (2022). Caries detection on intraoral images using artificial intelligence. Journal of Dental Research, 101(2), 158–165. https://doi.org/10.1177/00220345211032524

[20] Liang, Y., Li, D., Deng, D., Chu, C. H., Mei, M. L., Li, Y., Yu, N., He, J., & Cheng, L. (2025). AI-driven dental caries management strategies: From clinical practice to professional education and public self care. International Dental Journal, 75(2), 100827. https://doi.org/10.1016/j.identj.2025.04.007

[21] Liu, Y., Cheng, Y., Song, Y., Cai, D., & Zhang, N. (2024). Oral screening of dental calculus, gingivitis and dental caries through segmentation on intraoral photographic images using deep learning. BMC Oral Health, 24, 1287. https://doi.org/10.1186/s12903-024-05072-1

[22] Luke, A. M., & Rezallah, N. N. F. (2025). Accuracy of artificial intelligence in caries detection: A systematic review and meta-analysis. Head & Face Medicine, 21, 24. https://doi.org/10.1186/s13005-025-00496-8

[23] Marwaha, J., Singla, M., Nath, A., & Arya, A. (2025). Revolutionizing the diagnosis of dental caries using artificial intelligence-based methods. Journal of Conservative Dentistry and Endodontics, 28(5), 401–405. https://doi.org/10.4103/JCDE.JCDE_172_25

[24] Moharrami, M., Farmer, J., Singhal, S., Watson, E., Glogauer, M., Johnson, A. E. W., Schwendicke, F., & Quinonez, C. (2024). Detecting dental caries on oral photographs using artificial intelligence: A systematic review. Oral Diseases, 30(4), 1765–1783. https://doi.org/10.1111/odi.14659

[25] Negi, S., Mathur, A., Tripathy, S., Mehta, V., Snigdha, N. T., Adil, A. H., & Karobari, M. I. (2024). Artificial intelligence in dental caries diagnosis and detection: An umbrella review. Clinical and Experimental Dental Research, 10, e70004. https://doi.org/10.1002/cre2.70004

[26] Prados-Privado, M., García Villalón, J., Martínez-Martínez, C. H., Ivorra, C., & Prados-Frutos, J. C. (2020). Dental caries diagnosis and detection using neural networks: A systematic review. Journal of Clinical Medicine, 9(11), 3579. https://doi.org/10.3390/jcm9113579

[27] Qayyum, A., Tahir, A., Butt, M. A., Luke, A., Abbas, H. T., Qadir, J., Arshad, K., Assaleh, K., Imran, M. A., & Abbasi, Q. H. (2023). Dental caries detection using a semi-supervised learning approach. Scientific Reports, 13, 749. https://doi.org/10.1038/s41598-023-27808-9

[28] Ramezanzade, S., Dascalu, T. L., Ibragimov, B., Bakhshandeh, A., & Bjørndal, L. (2023). Prediction of pulp exposure before caries excavation using artificial intelligence: Deep learning-based image data versus standard dental radiographs. Journal of Dentistry, 138, 104732. https://doi.org/10.1016/j.jdent.2023.104732

[29] Schropp, L., Sørensen, A. P. S., Devlin, H., & Matzen, L. H. (2024). Use of artificial intelligence software in dental education: A study on assisted proximal caries assessment in bitewing radiographs. European Journal of Dental Education, 28(2), 490–496. https://doi.org/10.1111/eje.12973

[30] Schwendicke, F., Cejudo Grano de Oro, J., Garcia Cantu, A., Meyer-Lueckel, H., Chaurasia, A., & Krois, J. (2022). Artificial Intelligence for Caries Detection: Value of Data and Information. Journal of dental research, 101(11), 1350–1356. https://doi.org/10.1177/00220345221113756

[31] Schwendicke, F., Rossi, J. G., Göstemeyer, G., Elhennawy, K., Cantu, A. G., Gaudin, R., Chaurasia, A., Gehrung, S., & Krois, J. (2021). Cost-effectiveness of artificial

intelligence for proximal caries detection. Journal of Dental Research, 100(4), 369–376. https://doi.org/10.1177/0022034520972335

[32] Schwarzmaier, J., Frenkel, E., Neumayr, J., Ammar, N., Kessler, A., Schwendicke, F., Kühnisch, J., & Dujic, H. (2024). Validation of an artificial intelligence-based model for early childhood caries detection in dental photographs. Journal of Clinical Medicine, 13(17), 5215. https://doi.org/10.3390/jcm13175215

[33] Szabó, V., Szabó, B. T., Orhan, K., Veres, D. S., Manulis, D., Ezhov, M., & Sanders, A. (2024). Validation of artificial intelligence application for dental caries diagnosis on intraoral bitewing and periapical radiographs. Journal of Dentistry, 147, 105105. https://doi.org/10.1016/j.jdent.2024.105105

[34] Vinayahalingam, S., Kempers, S., Limon, L., Deibel, D., Maal, T., Hanisch, M., Bergé, S., & Xi, T. (2021). Classification of caries in third molars on panoramic radiographs using deep learning. Scientific Reports, 11, 12609. https://doi.org/10.1038/s41598-021-92121-2

[35] Xiong, Y., Zhang, H., Zhou, S., Lu, M., Huang, J., Huang, Q., Huang, B., & Ding, J. (2024). Simultaneous detection of dental caries and fissure sealant in intraoral photos by deep learning: A pilot study. BMC Oral Health, 24, 553. https://doi.org/10.1186/s12903-024-04254-1

[36] Zhang, J.-W., Fan, J., Zhao, F.-B., Ma, B., Shen, X.-Q., & Geng, Y.-M. (2024). Diagnostic accuracy of artificial intelligence–assisted caries detection: A clinical evaluation. BMC Oral Health, 24, 1095. https://doi.org/10.1186/s12903-024-04847-w

[37] Duangthip, D., & Chu, C. H. (2020). Challenges in Oral Hygiene and Oral Health Policy. Frontiers in oral health, 1, 575428. https://doi.org/10.3389/froh.2020.575428

Strength vs. Unity: A Comparative Analysis of Nationalist Rhetoric in Trump and Xi Jinping's Foreign Policy Speeches By David Meyers

Abstract

An often-overlooked factor influencing the foreign policy decisions of the United States and the People's Republic of China is the impact that foreign policy rhetoric has on fostering domestic legitimacy for each administration. While nationalist strategies used by both countries have been documented separately in scholarly literature, no prior study has directly compared the Trump and Xi administrations' rhetoric using a shared coding framework. Due to this, the research question pursued was: in what ways did the Trump administration (2017-2021) and the Xi Jinping administration (post-2016) employ nationalistic foreign policy rhetoric to maintain domestic legitimacy? To answer this, four primary source speeches concerning U.S.-China relations were assembled, two delivered by officials within the Trump administration and two by Xi Jinping. These speeches were analyzed using open and then axial coding across six categories of nationalist rhetoric: threat framing, claims of national superiority, calls for unity or sacrifice, anti-foreigner or anti-globalization messaging, sovereignty affirmations, and appeals to historical or cultural context. It was found that the Trump administration's speeches relied heavily on threat framing, with 24 instances present, portraying China as an antagonistic entity in order to justify a combative posture to the American public. In contrast, the Xi speeches did not contain a single instance of threat framing, and instead relied most heavily on appeals to shared historical and cultural context, with 17 instances present, in order to promote unity among the Chinese people. This suggests that the difference in regime type between the two nations may be the reason for the differing forms nationalist rhetoric took within each administration, rather than whether it was present at all. Due to the study's small sample size and single researcher, these findings should be read as suggestive of patterns rather than as generalizable claims about each administration as a whole.

Introduction

The current geopolitical reality of the world's political sphere is defined by two global powers, the United States of America and the People's Republic of China. Both of these countries' foreign policy decisions are among the most impactful in the world, affecting the lives of billions of people. However, an often-overlooked factor influencing both countries' foreign policy decisions is the impact of specific foreign policy actions in fostering domestic legitimacy for the ruling administration (Magcamit 7). The inquiry into this topic is essential to the understanding of the root causes of the foreign policy actions of the two most powerful nations in the world, with the two nations having a combined population of 1.749 billion people and producing 43 percent of the world's nominal GDP. These decisions impact people and states far beyond the borders of the United States and China, and understanding the causes of the respective countries' foreign policy decisions would aid in predicting how they would affect these billions (International Monetary Fund).

While both of the approaches the United States and China have taken to secure domestic legitimacy have been extensively documented in scholarly literature as of the present day, analysis of the comparative aspect of the two approaches is lacking in the current body of scholarly research. Therefore, there exists a gap in the comparative understanding of the two methods that must be addressed. Understanding the comparative methods for fostering domestic legitimacy through foreign policy and the reasons why both countries perform these actions is essential for understanding why each country performs certain foreign policy actions and how both the United States and China can be expected to act on the world stage. Being that the geopolitical rivalry between the two countries has been present before and throughout the administrations that the study focused on, understanding the motivations behind the United States' and China's use of nationalist rhetoric is important beyond a certain time period. No prior study has explicitly and directly compared the Trump and Xi's administration's nationalist rhetoric using a shared coding framework. Due to this, the question that the research pursued was: In what ways did the Trump administration (2017-2021) and the Xi Jinping administration (post-2016) employ nationalistic foreign policy rhetoric to maintain domestic legitimacy? The research answered this question through comparative analysis of multiple primary source documents and speeches given by representatives of the administrations of both countries.

Literature Review

Pre-Trump and Xi Background

In recent years on the world stage, there has been a significant increase in nationalist rhetoric from the two premier powers of the world, the United States of America and the People's Republic of China. Nationalist rhetoric is defined by patriotic statements about the quality and virtues of the country, often by contrasting those virtues against stated flaws of other nations (Sperling 513). In the United States, this is not a new phenomenon. Due to the United States being the unquestionably most powerful nation in the world, their foreign policy outlook has always been naturally hegemonic, with the country understandably attempting to sustain its position as the world's dominant power (Mendes 7). Many experts agree that this hegemony leads to a nationalistic outlook by both the leaders of the nation and the people themselves, as the hegemonic position is now not just a geopolitical and foreign policy issue for the United States public, but also an issue of legitimacy, as by and large, the people of the United States seemed to have taken the fact that their nation is a hegemon as an immutable fact of reality, and any threat to their hegemony would also constitute a threat to the legitimacy of the current government and of part of the American nationalist and exceptionalist identity (Carpenter 36).

In contrast, over the last several decades, the People's Republic of China (PRC) has become the primary challenger to the United States on the world stage, with China's share of world GDP increasing significantly, growing from 4 percent of the world's total GDP in 1990 to 22 percent of the total in 2022. The PRC has used this rapid economic growth to assert itself on the world stage, creating economic initiatives such as the Belt and Road Initiative, a network of

ports and infrastructure designed to benefit Chinese trade interests (Mochtak and Turcsanyi 745). The nationalism in current Chinese foreign policy stems from the historical strength of the Chinese state, with the current government of China embarking on a campaign of "national rejuvenation", a mission to restore China's "rightful" place as a leader on the world stage (Akdağ 206).

Ultimately, the current nationalist rhetoric and foreign policy actions of the two premier powers on Earth both stem from the historical background of their respective nations, and their efforts to either restore or maintain their geopolitical positions, as well as maintain legitimacy at home when making these claims (Magcamit 8).

U.S. – Populism and Foreign Policy under Trump (2016–2020 and beyond)

The nationalist rhetoric of the United States heightened significantly with the election of president Donald Trump in 2016. Immediately, it was clear that president Trump would take a much more populist approach to foreign policy than his predecessors, as evidence from him threatening and ultimately succeeding to pull out of the Paris Climate Accords in 2017, which according to him, was a "globalist" institution (Wojczewski 295). Trump also implemented protectionist trade policies in his first term, and again in his second term, attempting to foster growth in American industry, and making a concentrated effort to roll back America's connectivity with the rest of the global economy. The motivation behind these moves is debated by those within the field, but many see it as being due to Trump being an adherent to the nationalist "America First" foreign policy and economic policy outlook, which dictates that America should pull out of foreign affairs and move industry back from overseas (Butterfield).

However, this nationalist rhetoric isn't solely attributed to Trump's personal beliefs. Although Trump's nationalist foreign policy beliefs play a role in the American posturing under his administration, it is also done to foster legitimacy in the government at home. Trump's main appeal and legitimacy as a populist is that he positions himself as a man of the people, helping to protect the general population of America more than perceived globalist institutions (Wojczewski 298). In this regard, the nationalistic foreign policy outlook that may seem brazen and outrageous to some is actually done to appeal to Trump's nationalist and "America First"-minded voter base, making his government seem more legitimate (Patrick).

Despite this, Trump's first term seemed to have been reigned in by the American foreign policy establishment, as evidenced by his posturing about leaving NATO not coming to fruition. This is due to the fact that the United States foreign policy establishment that is made up of defense contractors and other economic special interests ensures that each president's administration does not deviate enough from policy that would benefit them, even if it comes at a cost to that administration's domestic legitimacy (Schultz 34).

Understanding the specific variety of Trump's nationalist rhetoric within his first term and comparing it to that utilized by Xi Jinping could provide key insights into the deeper motivations of not just the two administrations, but the larger foreign policy establishments in the United States and China respectively.

China – Populism and Foreign Policy under Xi Jinping (2013-present)

In the People's Republic of China, the approach of using nationalist foreign policy to secure domestic legitimacy has a few notable differences from the approaches used by the United States of America. Firstly, the Chinese public opinion towards their foreign policy is being skewed by the lens of them being a rising power challenging a ruling power, which is the United States and their allies (Zhai 73). This means that the type of foreign policy that the PRC will exhibit in order to maintain domestic legitimacy will be primarily nationalist, as well as antagonistic in regards to the established "ruling power": the United States (Zhai 74).

A type of nationalistic diplomacy that the PRC uses is known as "Wolf Warrior Diplomacy", which consists of having diplomats say intentionally inflammatory statements towards countries that are perceived to be the enemy (such as the United States or Japan) so that the population at home gets riled up with nationalist fervor (Kong). Wolf Warrior diplomacy has been employed by the PRC on a number of occasions during Xi Jinping's administration, including during COVID-19, a time of crisis for the regime, as they needed to justify their (while effective) heavy handed approach to stopping the virus. High-ranking officials in the Chinese Ministry of Foreign Affairs such as Zhao Lijian, posted provocative things on social media, including a tweet containing an artist's interpretation of Australian war crime allegations (Zhang et al.). These tweets were intentionally provocative to stoke nationalist fervor at home and to aid the Chinese government in maintaining its legitimacy in regards to its COVID policy. The Chinese government also heavily employs inflammatory rhetoric concerning the status of Taiwan, insisting that the island is an integral part of the Chinese homeland, and often threatening invasion.

The reason that the PRC is utilizing foreign policy and diplomacy so heavily as a way to maintain domestic legitimacy stems from the fact that they are an authoritarian country, so they don't have the built-in legitimacy that many democracies, such as the United States, do. That means that the PRC and other authoritarian countries rely on other means to maintain their government's legitimacy, such as the PRC's previous plan of maintaining high economic growth, similar to Singapore's strategy. However, the PRC's economic growth has slowed significantly, and their quality of life has stagnated in recent years, so they made the choice to pivot to foreign policy as a means of maintaining domestic legitimacy (Congiu and Uras 22).

While it is known that China and the United States have been and are currently at odds with each other on the world stage, and that both the Trump and Xi administrations have utilized nationalist rhetoric in attempts to foster domestic legitimacy, it is still not yet clear how each administration's use of nationalist rhetoric compares to the other. This is of key importance to understanding the deeper motivations behind the attempts of both administrations to utilize this nationalist rhetoric in order to increase their own domestic legitimacy, and understanding what specific foreign policy actions each administration wishes to be seen as legitimate in.

Method

The method of the research's case study attempted to address the question: In what ways did the Trump administration (2017-2021) and the Xi Jinping administration (post-2016) employ nationalistic foreign policy rhetoric to maintain domestic legitimacy? The researcher used a qualitative approach to analyzing various documents that would contain the information needed to answer this research question. The researcher collected information from major US and Chinese speeches during the time period about unfriendly nations and addressing their foreign policy in regards to those countries. All of these sources were needed to answer the research question as both the US and Chinese governments use these avenues of communication to reach as many people as possible, which shows that the foreign policy nationalism presented in those sources is the foreign policy that the respective governments of the United States and China during the selected administrations wished to broadcast to their populace. The documents were not solely sourced from Trump and Xi themselves, but rather also included the statements of the high-level diplomats part of each administration, as this yielded more relevant sources to pull from. This was modeled after the 2002 study conducted by Hongying Wang, as his approach to tracking keyword and phrase frequency allowed for accurate tracking of the differences between the nationalistic rhetoric exhibited by Chinese administrations (Wang). This case study method was appropriate because it allows the comparative analysis gap to be filled most effectively, as actual speeches and rhetoric used by both parties can be analyzed rather than the reaction to said rhetoric, which could obscure the accuracy of the data due to the political bias inherent within the US/China relations issue.

Step 1

First, the study assembled four sources during the time period of each administration that addressed the subject matter of foreign policy of interacting with unfriendly countries, specifically about the relations between the United States and China. Two speeches from each country's administration were selected to be analyzed, numbering four in total. The speeches used for the United States under the Trump administration were the remarks delivered by Vice President Mike Pence on October 4th, 2018 (O'Brien 11), and the remarks given by US President Donald Trump on May 30th, 2020 (O'Brien 52). The speeches used for the People's Republic of China under the Xi administration were Xi's speech at the welcome dinner by Friendly Organizations in the United States (Jinping), as well as the speech by H.E. Xi Jinping President of the People's Republic of China at the Conference Marking the 50th Anniversary of the Restoration of the Lawful Seat of the People's Republic of China in the United Nations (Jinping). These four speeches were all regarding the relationship between the United States and China, as well as the relationship between each of those countries and the rest of the world. Each of these speeches provide detailed information about how each country publicly views the other, and both of them contained enough nationalistic rhetoric to perform a thorough comparative analysis.

Step 2

Next, each of these documents were analyzed to track the instances of six different categories of nationalistic rhetoric throughout the content of these documents. These categories were: “Us vs Them” or threat framing language, claims of national superiority, calls for unity or sacrifice, anti-globalization or anti-foreigner messaging, sovereignty affirmations, and appeals to broader historical or cultural context. Each of these categories encompass specific types of commonly used nationalistic rhetoric, which made them a valuable lens in which to code and view the speeches that were studied.

Threat Framing Language	-Referring to another country or entity as being in opposition or being harmful to the home country
Claims of national superiority/national achievement	-Claiming that the home country is superior to other countries, or that the achievements of the home country are exceptional. These claims can be falsifiable or non-falsifiable.
Calls for unity or sacrifice	Rhetoric that emphasizes collective struggle or togetherness as a country
Anti-foreigner/Anti-Globalization messaging	-Messaging that frames an unfriendly country or international entity in a negative context.
Sovereignty affirmations	-Rhetoric that primarily appeals to the independence of the country to make their own decisions regarding their own affairs
Appeals to broader historical or cultural context	-Making references to historical events or the culture of the country

Step 3

After assembling the sources, each of them were analyzed six times, with each reading focusing on a different category of analysis. The study first analyzed each text for instances of threat framing, color coding each instance of it as it appeared in the paper. The study utilized a number of key words and phrases to identify threat framing during coding analysis and reading, then analyzed each category more thoroughly to ensure that the sentence truly fit the experimental definition of threat framing. This process was repeated for each category of nationalist rhetoric. Each instance of the certain categories of nationalist rhetoric were recorded in a data table, with each instance having a quote and citation. The table also recorded the total number of instances of each type of nationalist rhetoric for each document used. The process used to determine whether a certain phrase in the documents correlated to one of the coding categories was open coding. Each of the documents was broken into segments and labeled with color codes to identify segments and patterns while in the text, as demonstrated using the example of threat framing. The process of axial coding was subsequently used to group similar ideas found in the documents during the first round of coding into similar categories and themes (Leedy and Ormrod 297). After the data was analyzed multiple times, themes were developed in order to obtain a general sense of patterns within the data and to place the data into the

referenced themes and categories. The final step was to describe and interpret the data for further analysis. The data was thoroughly examined multiple times in order to triangulate similarities, differences, and irregularities. The coded data demonstrated the frequency in which each country, the United States and China, utilized various types of nationalistic rhetoric in specific contexts. This data demonstrated the various ways in which China and the United States under the specific administrations used nationalistic foreign policy rhetoric to foster domestic legitimacy, as the types of nationalistic rhetoric in foreign policy exhibited by each administration can be used as a baseline for inferring what kind of domestic legitimacy each wished to foster. The coded data was reviewed multiple times to ensure reliability, validity, and accuracy, as well as being checked for miscategorized items and phrases. Doing this also aided the coding process to be objective and aligned with the established definitions. No institutional review board approval was required for this study, as all of the data being used for the study was publicly available by the governments of both countries.

Results

For the comparative analysis of each document, the number of instances of each theme appearing in the documents, those being the speeches, were tallied up and assembled into three charts. One table provides the number of instances per theme for the Xi Jinping speeches alongside examples for each of the themes. The second table does the same for the Trump administration speeches that were coded and analyzed, and the third table compares the frequency of themes between both the Trump and the Xi speeches, displaying the difference in distribution between the two administrations in the frequency of the themes that were present. The differing distribution of themes as shown in the tables allows for specific trends in each administration's rhetoric to be understood.

Theme	Number of Instances	Example
Threat Framing	0	No example
Claims of National Superiority/achievement	5	"China is both a super-large economy and a super-large market. Not long ago the sixth China International Import Expo was held, attracting over 3,400 business exhibitors from 128 countries including the United States. The exhibition area of American companies has been the largest for six consecutive years at the Expo. Modernization for 1.4 billion Chinese is a huge opportunity that China provides to the world."
Calls for unity or sacrifice	7	"Through much hard work, the Chinese people have attained the goal of fully building a moderately prosperous society on the vast land of China, and won the battle against poverty, thus securing a historic success in eradicating absolute poverty. We have now embarked on a new journey toward fully building a modern socialist country and opened up bright prospects for the great rejuvenation of the Chinese nation."

Anti-foreigner/Anti-Globalization messaging	0	No example
Sovereignty affirmations	4	“Likewise, the United States should not bet against China, or interfere in China's internal affairs. It should instead welcome a peaceful, stable and prosperous China.”
Appeals to broader historical or cultural context	17	“We are committed to peaceful development to build a community with a shared future for mankind. Peace, amity and harmony are values embedded in Chinese civilization. Aggression and expansion are not in our genes.”

Fig 1: Xi Jinping (China) Coding Results

Figure 1 demonstrates the frequency of the presence of each theme within the Xi Jinping speeches. As shown, the most common themes present in the Xi speeches, in order, were: Appeals to shared culture/history (17), Calls for national unity, cooperation, and/or sacrifice (7), claims of national superiority and/or achievement relative to other countries (5), and sovereignty affirmations (4). Interestingly, the Xi speeches did not have a single instance of threat framing or anti-foreigner/anti-globalization messaging. The tone of the Xi Jinping speeches is much more history-focused and broadly unifying, as demonstrated by the appeals to broader history and culture and calls for unity, cooperation, and sacrifice being first and second in frequency respectively. As shown in many of the quotes provided and in the frequency distribution of the themes, Xi Jinping’s nationalist rhetoric primarily extols the virtues of China, Chinese civilization, and the Chinese people as a whole rather than being combative or threatening to other countries. As shown by the quote provided for the appealing to broader historical and cultural context theme, Xi portrays China and Chinese civilization as being inherently good and nonviolent by insisting that “Peace, amity and harmony are values embedded in Chinese civilization. Aggression and expansion are not in our genes.” Keeping with the non-confrontational approach to the administration’s nationalistic rhetoric in their speeches, the instances of sovereignty affirmations within the Xi speeches were less overtly aggressive and assertive, as shown by Xi stating that “Likewise, the United States should not bet against China, or interfere in China's internal affairs. It should instead welcome a peaceful, stable and prosperous China.” These quotes demonstrate that the most common nationalist rhetoric exhibited by the Xi administration in this context is in the form of positive affirmations and statements about China and the Chinese people’s characteristics, and not any overtly hostile statements.

Theme	Number of Instances	Example

Threat Framing	24	"China raided our factories, offshored our jobs, gutted our industries, stole our intellectual property, and violated their commitments under the World Trade Organization."
Claims of National Superiority/achievement	3	"Under President Trump's leadership, the United States of America has been defending our interests with renewed American strength."
Calls for unity or sacrifice	3	"Fortunately, Americans aren't buying it. For example, American farmers are standing with this President and are seeing real results from the strong stands that he's taken…"
Anti-foreigner/Anti-Globalization messaging	2	"Beijing now requires many American businesses to hand over their trade secrets as the cost of doing business in China. It also coordinates and sponsors the acquisition of American firms to gain ownership of their creations. Worst of all, Chinese security agencies have masterminded the wholesale theft of American technology—including cutting-edge military blueprints. And using that stolen technology, the Chinese Communist Party is turning plowshares into swords on a massive scale."
Sovereignty affirmations	7	"Despite such reckless harassment, the United States Navy will continue to fly, sail, and operate wherever international law allows and our national interests demand. We will not be intimidated and we will not stand down."
Appeals to broader historical or cultural context	6	"Under President Trump's leadership, the United States has taken decisive action to respond to China with American action, applying the principles and the policies long advocated in these halls."

Fig 2: Trump administration (USA) speech coding results

Figure 2 demonstrates the frequency of the presence of each theme within the Trump administration speeches. As shown above, the most common themes present in the Trump administration Speeches, in order, were: Threat Framing (24), Sovereignty affirmations (7), Appeals to national identity and culture (6), Claims of national superiority/accomplishment (3), Calls for unity/sacrifice (3), and anti foreigner/anti-globalization (2). The tone of the Trump speeches is much more combative and directly patriotic, as shown by the 24 instances of threat framing. This is also demonstrated by sovereignty affirmations and appeals to broader historical and cultural context being second and third in frequency. The distribution in theme frequency and the quotes provided both indicate that the Trump administration's nationalistic rhetoric on this topic is primarily combative against China and assertive about American capabilities and sovereignty. As demonstrated in the Trump administration quote for threat framing in the table, China is portrayed as being antagonistic or even hostile to the United States, its economy, its government, and its citizens. China is portrayed as a hostile entity in the administration's nationalistic rhetoric by insisting that "China raided our [USA's] factories, offshored our jobs,

gutted our industries, stole our intellectual property, and violated their commitments under the World Trade Organization." Continuing with the confrontational rhetoric portraying China as an entity hostile to the United States, the Trump administration speeches used anti-foreigner and anti-China messaging, stating that "...Chinese security agencies have masterminded the wholesale theft of American technology—including cutting-edge military blueprints. And using that stolen technology, the Chinese Communist Party is turning plowshares into swords on a massive scale." This demonstrates that the Trump administration's nationalist rhetoric in this context is primarily combative and hostile. While the Trump administration speeches do use nationalist rhetoric consisting of positive affirmations and statements about the United States, the nationalist rhetoric in these speeches are confrontational, combative, and sometimes hostile.

Instances of each theme per set of speeches	Threat Framing	Claims of National Superiority/achievement	Calls for unity or sacrifice	Anti-foreigner/Anti-Globalization messaging	Sovereignty affirmations	Appeals to broader historical or cultural context
China (Xi)	0	5	7	0	4	17
USA (Trump)	24	3	3	2	7	6

Fig 3: Comparative Result Frequency

Figure 3 demonstrates the distribution of the frequency of themes between the Trump and Xi speeches. The Xi speeches tended to have a tone of cooperation and unity overall, which is in contrast to the speeches made by the Trump administration, which were more combative. Compared to the speeches made by the Xi administration, the Trump speeches focused much less on appeals to broader historical and cultural context, with only 6 instances compared to the 17 from Xi's speeches. Both the Xi and Trump speeches had different levels of emphasis on national unity, cooperation, and sacrifice, with 3 instances from Trump and 7 from Xi. Xi's speeches had a higher level of emphasis on national unity, cooperation, and sacrifice, urging national unity and cooperation when utilizing nationalist rhetoric much more than Trump's speeches did. The Xi administration's nationalist rhetoric largely focused on advocating for the virtues of China with very few combative instances, while the Trump administration's nationalist rhetoric, while still advocating for the virtues of the United States, was much more balanced by combative rhetoric towards China.

Discussion

Both of the Xi and the Trump speeches had different focuses and manners in which they used nationalist rhetoric. The Trump speeches were especially focused on asserting American strength and resolve, as shown by his administration's comments about the US Navy with

regards to how it deals with China, where Vice President Mike Pence states "Despite such reckless harassment, the United States Navy will continue to fly, sail, and operate wherever international law allows and our national interests demand. We will not be intimidated and we will not stand down." This demonstrates that the Trump speeches, by discussing American strength and resolve in a generally positive light, which is patriotic and nationalist rhetoric, attempts to increase the willingness of the home country's population to support the cause by associating the cause of opposing China with American values.

The Xi speeches specifically emphasize traditional Chinese values and civilization and frame them in terms of a unifying force, stating them as values and history all Chinese people share. Xi states this about the Chinese people: "We are committed to peaceful development to build a community with a shared future for mankind. Peace, amity and harmony are values embedded in Chinese civilization. Aggression and expansion are not in our genes." This demonstrates that in the Xi speeches, the Chinese values that are being mentioned are deliberately used to prove a point about the character of the Chinese people and of the nation, which is that they are peaceful and harmonious. This is nationalistic rhetoric, as it asserts that China is a peaceful and harmonious country, two words that have positive connotations. The Xi speeches tended to have a tone of cooperation and unity overall, which is in contrast to the speeches made by the Trump administration, which were more combative. In the Xi speeches, there were no instances of threat framing present, while there were 24 instances present in the Trump speeches. An example of the Xi speeches demonstrating cooperation and unity is "For these 50 years, the Chinese people have stood in solidarity and cooperation with people around the world and upheld international equity and justice, contributing significantly to world peace and development. The Chinese people are peace-loving people and know well the value of peace and stability." This stands in sharp contrast to the combative tone of the Trump speeches, which include instances of threat framing such as "But our message to China's rulers is this: This President will not back down. The American people will not be swayed. And we will continue to stand strong for our security and our economy, even as we hope for improved relations with Beijing." This demonstrates that the strategies that the Xi and Trump speeches use utilize different manners of attempting to use nationalism. Xi's speeches use language of cooperation and respect in order to portray China, the Chinese government and the Chinese people in a positive light, which is patriotic and somewhat nationalistic rhetoric.

The Trump speeches utilize nationalist rhetoric, such as portraying other countries in a combative or threatening light, in order to create a sense of American unity. The Trump speeches, with their significantly antagonistic tone against China, especially in regards to how China's actions hurt Americans and how the US government is trying to protect them, also attempt to increase an American sense of opposition towards China specifically. This demonstrates that the Trump administration in this case from portraying China in a bad or combative light through this kind of nationalistic rhetoric is to convince the population that China is the enemy, and make sure that their administration is the one to be trusted to stand up to this enemy. The Xi speeches had an increased emphasis on national unity, cooperation, and

sacrifice than the Trump speeches did, with 3 instances from Trump and 7 from Xi. This demonstrates that the Xi speeches and administration are trying to appeal much more to the collective of China and their struggle to call them to action than the Trump speeches are attempting to appeal to the collective people of the United States. The Xi speeches were especially trying to create a sense of Chinese solidarity and collective struggle, repeatedly mentioning and appealing to Chinese history and culture to create a sense of national unity and solidarity.

In the Xi speeches, there were 17 instances of appealing to shared Chinese historical and cultural context, including "The path of socialism with Chinese characteristics has been found under the guidance of the theory of scientific socialism, and is rooted in the tradition of the Chinese civilization with an uninterrupted history of more than 5,000 years. We are proud of our choice, just as you are proud of yours." Xi consistently uses these appeals in order to create a sense of national unity and solidarity among the Chinese people, which is a method of using nationalist rhetoric. Using rhetoric that appeals to shared cultural background and history is used to strengthen the notion of a shared cultural identity and values, which Xi associates with the Chinese government that he leads.

Overall, through the analysis of this data, it was surmised that the speeches during the Trump administration utilized a significant amount of combative and confrontational nationalist rhetoric in order to foster domestic legitimacy for said administration. In contrast, the Xi speeches tended to be less combative and confrontational, and the nationalist rhetoric in those speeches took the form of appealing to shared historical values and culture in order to increase the administration's domestic legitimacy.

Limitations and Implications for Future Research

There were several limitations that were present within this study, one being the sample size of the case study. There were only two speeches for each administration being studied, and they were chosen regardless of context other than the required topic of US-China relations. Because of this, results may vary if speeches were chosen from different time periods and other phases of US-China relations. For this reason, it should be noted that the results of this study are only applicable to the specific time period and context of each of the speeches. Due to the sample size, the findings in this study should be read as suggestive of patterns within this specific sample and not as generalizable claims to each administration as a whole. If future research were to be conducted on this topic, the sample size of the study should be expanded significantly to take into account multiple different conditions of US-China relations and explore how the nationalist rhetoric used to increase domestic legitimacy on both sides shifts depending on the affairs and context of the time. An additional limitation of this study is only having one researcher, meaning that there was no inter-rater reliability. Only having one researcher, especially for a case study analysis, increases the risk of having personal bias interfere slightly with the interpretation of the results. Future research could utilize more researchers to minimize personal bias interfering when choosing the speeches and interpreting the coded data.

There were also several implications that resulted from this study. One major implication concerns the tone of the Trump administration speeches against China. The Trump speeches had a significant combative tone against China, especially in regards to how China's actions hurt Americans and how the US government is trying to protect them. This would be of significant importance to both American and Chinese policymakers and international foreign policy analysts as it reflects the intent of the Trump administration when using nationalist rhetoric. For foreign policy scholars and international policy analysts and practitioners, the comparison suggests regime type (democracy vs authoritarianism) may shape the form and function of nationalist rhetoric more than whether nationalism is present at all. The rhetoric used by the Trump administration as analyzed in the study demonstrates that Trump and his administration are attempting to use combative and nationalist rhetoric to make their administration seem more legitimate, as the administration is stating that they are the ones protecting the people from harmful outside forces. Trump, being a president in a democratic country needing to win elections, possibly has his administration use this foreign policy rhetoric in order to increase voter support for him or the rest of the Republican party. This is in contrast to the less combative and confrontational tone of the Xi speeches. Future research could focus on what specific situations the Trump administration utilizes nationalist rhetoric in a combative tone, and which specific foreign policy actions they take before and afterwards. Another implication concerns the specific rhetoric of the Xi speeches in regards to shared Chinese history and cultural background. In the Xi speeches, the administration repeatedly mentions and appeals to Chinese history and culture to create a sense of national unity and solidarity. This would be important to foreign policy analysts and policymakers as it reflects the intent of the Xi administration when using nationalistic rhetoric. Xi is attempting to enforce and strengthen national unity and solidarity, as well as attempting to link his administration to the long history of Chinese civilization. Future research could focus on this specific aspect of the administration's speeches in more detail, especially which aspects of Xi's administration and the current Chinese government he tries to link with Chinese history, and which Chinese history the administration attempts to appeal to. This would reveal which parts of Chinese history the Chinese Communist Party finds important enough to link to in order to gain legitimacy.

The use of specific kinds of nationalist rhetoric in the foreign policy-related speeches of both the United States under the Trump administration and the PRC under the Xi administration provided an important view into the ways both attempt to foster domestic legitimacy. It was found that the Trump administration speeches primarily used threat framing-based nationalistic rhetoric, while the Xi administration speeches primarily used nationalist rhetoric which consisted of appeals to historical and cultural context. This implied that the Trump and the Xi administrations, running different types of government (democracy vs authoritarian), had significantly different aims when utilizing nationalist rhetoric in order to maintain domestic legitimacy.

Works Cited

Akdağ, Zeynep. "China's Assertive Foreign Policy and Global Visions under Xi Jinping." Akademik Incelemeler Dergisi, vol. 19, no. 1, 2024, pp. 204–211. https://doi.org/10.17550/akademikincelemeler/1423366.

Butterfield, L. "The New Right, Trade, and US Foreign Policy." Stanford University, 5 June 2025, casi.stanford.edu/news/new-right-trade-and-us-foreign-policy-0. Accessed 30 Oct. 2025.

Carpenter, Ted Galen. "The Populist Surge and the Rebirth of Foreign Policy Nationalism." SAIS Review of International Affairs, vol. 37, no. 1, 2017, pp. 33–46. https://doi.org/10.1353/sais.2017.0003.

Congiu, Fabrizio, and Alessandro Uras. "China 2016: Defending the Legitimacy of the Party-State's Authority." Asia Maior, vol. 27, 2016, pp. 17–51.

"GDP, Current Prices Billions of U.S. Dollars." International Monetary Fund, 9 Feb. 2026, www.imf.org/external/datamapper/NGDPD@WEO/OEMDC/ADVEC/WEOWORLD. Accessed 20 Apr. 2026.

Jinping, Xi. "Speech by H.E. Xi Jinping President of the People's Republic of China at the Conference Marking the 50th Anniversary of the Restoration of the Lawful Seat of the People's Republic of China in the United Nations." Embassy of the People's Republic of China in the United States of America, 21 Oct. 2021, us.china-embassy.gov.cn/eng/zgyw/202110/t20211025_9982254.htm. Accessed 20 Apr. 2026.

Jinping, Xi. "Full Text of Xi's Speech at Welcome Dinner by Friendly Organizations in the United States." Englishnews.cn, 16 Nov. 2023, english.news.cn/20231116/e8430b6ab19146f4a15bbed8465530bc/c.html. Accessed 23 Apr. 2026.

Kong, Weiming. "Rampant Nationalism Is Undermining China's 'Three Warfares.'" The Diplomat, 23 Aug. 2024, thediplomat.com/2024/08/rampant-nationalism-is-undermining-chinas-three-warfares/. Accessed 30 Oct. 2025.

Leedy, Paul D., and Jeanne Ellis Ormrod. Practical Research: Planning and Design. 11th ed., Pearson, 2016.

Magcamit, Michael. "Explaining the Three-Way Linkage between Populism, Securitization, and Realist Foreign Policies: President Donald Trump and the Pursuit of 'America First' Doctrine." World Affairs, vol. 180, no. 3, 2017, pp. 6–35.

Mendes, Pedro Emanuel. "The Dynamics of Change in United States Foreign Policy: Contexts, Leadership, and Hegemonic Legitimacy." Social Sciences, vol. 12, no. 10, 2023, p. 560. https://doi.org/10.3390/socsci12100560.

Mochtak, Michal, and Richard Q. Turcsanyi. "Studying Chinese Foreign Policy Narratives: Introducing the Ministry of Foreign Affairs Press Conferences Corpus." Journal of

Chinese Political Science, vol. 26, no. 4, 2021, pp. 743–761. https://doi.org/10.1007/s11366-021-09762-3.

O'Brien, Robert C., editor. President Trump on China: Putting America First. Trumpwhitehousearchives.gov, 2 Nov. 2020, trumpwhitehouse.archives.gov/wp-content/uploads/2020/11/Trump-on-China-Putting-America-First.pdf. Accessed 2 Mar. 2026.

Patrick, Stewart. "Trump's Greenland and Panama Canal Threats Are a Throwback to an Old, Misguided Foreign Policy." Carnegie Endowment for International Peace, 7 Jan. 2025, carnegieendowment.org/emissary/2025/01/trump-greenland-panama-canal-monroe-doctrine-policy?lang=en. Accessed 30 Oct. 2025.

Schultz, David. "American Foreign Policy in the Age of Donald Trump." Lithuanian Annual Strategic Review, vol. 17, no. 1, 2019, pp. 11–34. https://doi.org/10.2478/lasr-2019-0001.

Sperling, James, and Mark Webber. "Trump's Foreign Policy and NATO: Exit and Voice." Review of International Studies, vol. 45, no. 3, 2019, pp. 511–526. https://doi.org/10.1017/s0260210519000123.

Wang, Hongying. "National Image Building and Chinese Foreign Policy." University of Leeds, 29 Aug. 2002, universityofleeds.github.io/philtaylorpapers/pmt/exhibits/2582/wang_h.pdf. Accessed 12 Nov. 2025.

Wojczewski, Thorsten. "Trump, Populism, and American Foreign Policy." Foreign Policy Analysis, vol. 16, no. 3, 2019, pp. 292–311. https://doi.org/10.1093/fpa/orz021.

Zhai, Yida. "Public Opinion in a Rising Power: National and International Orientations among the Chinese Public." Journal of East Asian Studies, vol. 23, no. 1, 2023, pp. 71–94. https://doi.org/10.1017/jea.2023.3.

Zhang, Anny, et al. "Tweet Storm Shows China Aims to Project Power through Provocation." The Strategist, 9 Dec. 2020, www.aspistrategist.org.au/tweet-storm-shows-china-aims-to-project-power-through-provocation/. Accessed 30 Oct. 2025.

A Structural Analysis: How Incentives Shape the Value-Action Gap of Climate Change

By Nikki Samios

Climate change is among the most widely recognized global crises of the twenty-first century. Scientific consensus affirms that it is largely human-caused, and public opinion surveys demonstrate that the majority of the world acknowledges both its existence and severity. Yet despite this awareness, efforts to mitigate climate change remain inconsistent and insufficient. This discrepancy between belief and behavior, commonly referred to as the value-action gap, poses a major challenge to sustainable progress. In order to understand why this discrepancy persists, this paper will first establish the value-action gap of climate change, explaining why it is presented by climate inaction. It will then examine this gap through the economic lens of the prisoner's dilemma, focusing on incentive structures and how countries respond. Finally, it will examine a case study of the European Union versus India and analyze how these countries interact within this dilemma. This paper will ultimately show how the persistence of climate inaction reflects a misalignment between immediate incentives for nations and global collective outcomes, illustrating how the prisoner's dilemma, compounded by temporal discounting, uncertainty, and externalities, makes voluntary cooperation strategically irrational.

Establishing the Global Value-Action Gap of Climate Change

A value-action gap is a way to describe the disconnect between people's stated values and how they actually behave.[1] In many areas of life, individuals articulate strong beliefs but fail to consistently act on them. For example, someone may value a healthy lifestyle but may choose convenience over nutrition in their daily routine. This gap does not necessarily reflect hypocrisy or a lack of concern; rather, it often is a reflection of competing incentives and habits that make it difficult to align behavior with values. In the context of climate change, this concept becomes especially significant. While many recognize environmental issues and express concern about consequences, their everyday actions do not always align with these beliefs. This same concept extends beyond individuals to nations, where governments commit to addressing climate change but often fail to implement and enforce sufficient policies. Understanding how this gap originates and why it persists is essential to understanding why widespread awareness of climate change has not turned into equally widespread action. To highlight the extent of this disconnect, it is useful to first examine public attitudes towards climate change globally.

Global awareness of climate change is widespread and well-documented. The world's largest survey on climate change, "People's Climate Vote 2024," was conducted by the United Nations Development Programme in 2024.[2] This survey, representing statistically 87% of the world's population, concluded that climate change remains on the minds of the majority of

[1] Itamar Shatz, "The Value-Action Gap: Why People Don't Act in Accordance with Their Beliefs," *Effectiviology*, accessed April 19, 2026, https://effectiviology.com/value-action-gap/.

[2] United Nations Development Programme, "The World's Largest Survey on Climate Change is Out–Here's What the Results Show," June 27, 2024, https://climatepromise.undp.org/news-and-stories/worlds-largest-survey-climate-change-out-heres-what-results-show.

people, regardless of where they are from.[3] Around 80% of people around the world believe that governments should take stronger action to address climate change, with a majority expressing concern about its impacts on their lives.[4] Likewise, large shares of people across middle-income countries report that climate change is already affecting their communities, with many identifying it as a major threat to their wellbeing.[5] These findings suggest that awareness and concern are widespread across diverse economic levels and geography. However, despite this global recognition, emissions continue to rise, and large-scale policy shifts remain insufficient, as current national climate policy commitments still place the world on track for a temperature rise of 2.3-2.5°C by the end of the century.[6] Climate inaction cannot be explained by ignorance alone, but instead reflects a deeper systemic gap between what people believe and how they act. This inconsistency between values and actual behavior defines the value-action gap of climate change.[7] This concern, without corresponding action, therefore raises a complex but crucial question: why does this value-action gap exist? To answer this question, it is essential to analyze it through the lens of a broader framework. Examining international responses to climate change provides a useful benchmark for understanding these dynamics.

The Paris Agreement: An Attempt at International Cooperation

The Paris Agreement represents one of the most ambitious international efforts to mitigate climate change, bringing nearly every nation together to share the benefits of committing to reducing greenhouse gas emissions. In theory, a coordinated global effort such as this should result in meaningful progress, as the benefit of mitigation is shared. However, despite these shared benefits, cooperation remains inconsistent and insufficient. So, why does cooperation fail, even when all parties would benefit? How nations interpret uncertainty and perceive the impact of their actions is critical to understanding the value-action gap.

The Paris Agreement, adopted in 2015, established clear goals for reducing emissions and mitigating global warming, representing a clear collective global commitment.[8] Its primary objective is to keep the increase in the global average temperature below 2°C above pre-industrial levels and pursue efforts to cap warming at 1.5°C.[9] Rather than fixed mandates for each country, the Agreement relies on Nationally Determined Contributions (NDCs), in which

[3] United Nations Development Programme, "The World's Largest Survey on Climate Change."

[4] United Nations Development Programme, "The World's Largest Survey on Climate Change."

[5] Moira Fagan at el., "People in Middle-Aged Countries Say Climate Change is Affecting Their Community," Pew Research Center, November 5, 2025, https://www.pewresearch.org/global/2025/11/05/people-in-middle-income-countries-say-climate-change-is-affecting-their-community/

[6] UN Environment Programme, *Emissions Gap Report 2025* (Nairobi: UNEP, 2025), November 4, 2025, https://www.unep.org/resources/emissions-gap-report-2025.

[7] Guido Grunwald, Ali Kara, and John E. Spillan, "Involvement Matters: Navigating the Value-Action Gap in Business Students' Sustainability Transformation Expectations–A Cross-Country Kano Study," *Sustainability Science* 20, no. 6 (2025): 2035-2057, https://link.springer.com/article/10.1007/s11625-025-01681-z.

[8] United Nations Climate Change, "The Paris Agreement," accessed February 12, 2026, https://unfccc.int/process-and-meetings/the-paris-agreement.

[9] United Nations Climate Change, "The Paris Agreement."

countries set their own emissions reduction targets based on their circumstances.[10] These targets vary significantly across countries, as each country determines its own baseline year, targets, and level of ambition, resulting in a wide range of commitments from different countries that are not directly comparable. These contributions are updated every five years with the expectation that contributions will progressively increase over time.[11] This process, often referred to as a "ratchet mechanism," is designed to gradually strengthen global ambition by requiring countries to revise their commitments as time goes on.[12] This structure was designed to encourage widespread participation by allowing countries flexibility to set their own targets while still maintaining collective oversight. To support this system, the Agreement requires countries to regularly report their emissions and track their progress. These reports are reviewed through an international process, and every five years, a global assessment evaluates collective progress towards the long-term goals.[13]

This cooperative approach seems appealing in theory, and to some extent it has been successful. Since climate change mitigation produces benefits for all, even incremental cooperation can result in progress when widely implemented. The Paris Agreement has resulted in some measurable progress: projected warming trajectories for the end of the century have been reduced from over 3.5℃ to between 2.3℃ and 2.8℃ .[14] Additionally, the usage of renewable energy sources, such as wind, tidal, and particularly solar, has significantly increased in the past decade.[15] Many countries, regions, cities, and companies are establishing carbon neutrality targets, and zero-carbon solutions are becoming competitive across economic sectors, representing 25% of emissions.[16] It is projected that by 2030, zero-carbon solutions could be competitive in sectors that account for over 70% of global emissions.[17] These developments demonstrate how coordinated efforts can make some progress towards long-term emissions trajectories, even without clear enforcement measures.

However, despite this progress, outcomes have fallen significantly short of expectations, as current global emissions reductions are also insufficient to meet Paris Climate Agreement targets. Although nearly every nation has submitted NDCs, these pledges remain insufficient to meet the goals outlined in the Paris Agreement. According to the Climate Action Tracker, no countries are 1.5℃ compatible, meaning their current policies are not aligned with limiting

---

[10] United Nations Climate Change, "The Paris Agreement."

[11] "The Paris Agreement," United Nations, accessed February 12, 2026, https://www.un.org/en/climatechange/paris-agreement.

[12] Sophie Yeo, "Explainer: the 'Ratchet Mechanism' within the Paris Climate Deal," *Carbon Brief*, December 3, 2015, https://www.carbonbrief.org/explainer-the-ratchet-mechanism-within-the-paris-climate-deal/#:~:text=The%20ratchet%20mechanism%2C%20also%20known%20as%20the,how%20much%20they%20intend%20to%20reduce%20emissions.

[13] United Nations Climate Change, "The Paris Agreement."

[14] "6 things that have happened since the Paris Agreement," Think Landscape, December 9, 2025, https://thinklandscape.globallandscapesforum.org/100713/10-years-paris-agreement/.

[15] Think Landscape, "6 things that have happened since the Paris Agreement."

[16] United Nations Climate Change, "The Paris Agreement."

[17] United Nations Climate Change, "The Paris Agreement."

global warming to 1.5℃ above pre-industrial levels, and many remain categorized as critically and highly insufficient.[18] Numerous countries have also been either unwilling or unable to set appropriate targets to reduce emissions and effectively implement them.[19] Even if all current NDCs were met, the world would still be on track for a temperature rise of 2.3-2.5°C by the end of the century.[20] Under the current trajectory of existing policies (excluding fully implemented NDCs), the projected warming temperature would rise to 2.8°C.[21] To remain below the critical 1.5-degree threshold established in 2015, emissions must be cut by 42 percent overall by 2030 and 57 percent by 2035–reductions that are far beyond the current trajectory.[22]

Additionally, despite its framework and global support, there is no mechanism to enforce the commitments in the Paris Agreement.[23] Instead, the Agreement relies on transparency and voluntary compliance. The main formal consequence for a member failing to meet its targets is a meeting with a global committee of neutral researchers, which then works with struggling members to create new plans.[24] While this cooperative approach has encouraged some participation, it has also limited the ability to guarantee compliance. Since there are no consequences for failing to meet commitments, countries face little external pressure to adhere to the targets. This weak enforcement makes the Paris Agreement not one of obligation, but instead one that is voluntary, where compliance depends on alignment with a country's economic and political interests. While the Agreement encourages participation, it cannot ensure compliance. As a result, countries may be incentivized to set less ambitious goals or delay action, especially if they believe others will do the same. This creates a gap between stated commitments and actual outcomes, reinforcing the pattern of climate inaction despite widespread agreement on its importance.

This lack of enforcement is compounded by a continued reliance on fossil fuels, which remain dominant in global energy production. Despite rapid growth in renewable energy capacity over the past double decades, much of this growth has supplemented, but not replaced, high-emissions energy sources. As global economic growth drives an increase in energy demand, the challenge of meeting that demand with cleaner energy sources becomes even more difficult.[25] In many cases, fossil fuels remain the most accessible, affordable, and reliable option, especially for developing countries with fewer resources. As a result, despite some progress, dependence on

[18] Climate Action Tracker, "Countries," accessed April 21, 2026, https://climateactiontracker.org/countries/.

[19] Think Landscape, "6 things that have happened since the Paris Agreement."

[20] UN Environment Programme, *Emissions Gap Report 2025*.

[21] UN Environment Programme, *Emissions Gap Report 2025*.

[22] Li Cohen, "Global Climate Disaster Inevitable If Emissions Aren't Drastically Reduced by 2035, U.N. Warns," CBS News, October 25, 2024, https://www.cbsnews.com/news/climate-change-un-report-disaster-emissions-future/.

[23] Kathryn Tso, "How Are Countries Held Accountable under the Paris Agreement?," MIT Climate Portal, August 8, 2025, https://climate.mit.edu/ask-mit/how-are-countries-held-accountable-under-paris-agreement.

[24] Tso, "How Are Countries Held Accountable under the Paris Agreement?"

[25] Jason Hickel, *Less Is More: How Degrowth Will Save the World* (London: Windmill Books, 2021).

fossil fuels continues to limit the effectiveness of global efforts to address climate change. While the benefits of reducing emissions benefit all, the costs do not burden all.

This gap between collective goals and countries' actions reflects a deeper problem of coordination under uncertainty. Although all countries benefit from mitigation efforts, each must decide whether to bear the costs without knowing how others will act. This uncertainty creates a situation in which cooperation becomes dependent on expectations of compliance rather than assurances. Here, the incentives to delay action or rely on others seemingly outweigh the incentives for cooperation.

The Prisoner's Dilemma: Illustrating The Incentive Structure of Climate Inaction

At its core, climate change represents a large-scale collective action problem. The benefits of reducing emissions are shared globally, but the costs of mitigation are high. This creates strong incentives to free ride, where nations benefit from other nations' behavior without changing their own. Since the atmosphere is shared globally, no one is excluded from its benefits, and no one alone can stabilize it. As a result, countries are incentivized to delay climate action, relying on others to bear the burdens of mitigation while they benefit from others' contributions to climate stability.

This situation closely resembles a large-scale prisoner's dilemma, a fundamental concept of economic game theory that illustrates how rational decision-making can lead to collectively irrational outcomes. The prisoner's dilemma originates from a scenario in which two suspects are held in separate, isolated rooms, with no means of communication, and are questioned as to whether they committed a crime. Therefore, they must independently decide whether to deny or confess without knowing the other's choice. If both deny, and cooperate with each other, they achieve the best possible outcome. If one confesses while the other denies, the confessor receives a positive outcome while the denier receives a negative outcome. If both confess, they receive a worse outcome than if they held out and trusted each other not to confess. While mutual cooperation results in the best collective outcome, self-interest pushes players towards confession. Regardless of what the other does, confession is the dominant strategy, as it guarantees a better outcome for the player. This confession, a defection from cooperation, produces a worse result overall for both parties. In this case, what seems rational at the individual level creates a collectively irrational outcome.[26] Although climate change involves many more than two players, analyzing it through a simplified two-player prisoner's dilemma is useful for understanding the incentive structure.

While this model explains the basic structure of climate inaction, it does not fully reflect the reality of decision-making. In the context of climate change, countries are not basing their decisions on a single set of outcomes. Instead, they make decisions within two key frameworks: one relating to short-term economic costs, and another concerning long-term environmental

[26] CORE Econ, *The Economy*, accessed March 12, 2026, https://books.core-econ.org/the-economy-v1/en/.

impacts.[27] These conflicting incentive structures can be represented through two variations of the prisoner's dilemma. Figure 1 resembles a prisoner's dilemma model of short-term economic costs, while Figure 2 represents a model of long-term environmental damage. These models assume two countries, countries A and B, to be equal in terms of economy, infrastructure, and resources. In these models, each country is given two options: reduce emissions ("Take Action"), or do nothing ("Don't Act").

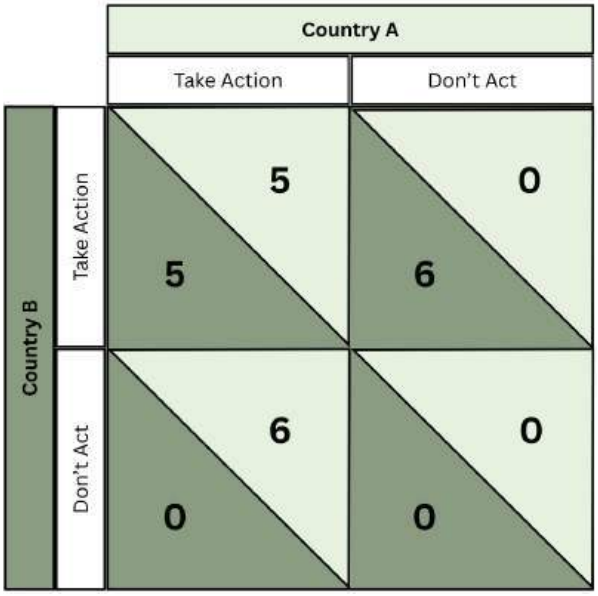


Figure 1

This model reflects short-term economic incentives, where greater numbers represent greater immediate costs associated with climate action. Taking action by reducing emissions requires mandates that will adversely impact economic growth, resulting in short-term costs for that country. In contrast, choosing inaction allows countries to maintain reliance on fossil fuels and prioritize economic growth, resulting in no economic costs for that country. If country A decides to reduce emissions while country B does not, country A will be worse off economically while country B will still benefit from country A's choice.[28] In this case, defection becomes the dominant strategy regardless of whether the other country takes action, since avoiding the immediate costs of action produces a better short-term outcome. This dynamic proves how climate policy often competes directly with economic priorities, creating a scenario in which inaction becomes the rational choice.

However, when the same decision is evaluated through a long-term environmental lens, the incentive structure significantly shifts. Figure 2 below illustrates the prisoner's dilemma

[27] Carrozo Magli Alessio, Piero Manfredi, "Coordination Games vs. Prisoner's Dilemma in Sustainability Games: A Critique of Recent Contributions and a Discussion of Policy Implications," *Ecological Economics* 192 (2022): 107268, https://www.sciencedirect.com/science/article/abs/pii/S092180092100327X#preview-section-abstract.

[28] Andreï-Bogdan Sterescu, "The Prisoner's Dilemma in Environmental Politics: One Model to Rule Them All?" *E-International Relations*, June 14, 2018, https://www.e-ir.info/2018/06/14/the-prisoners-dilemma-in-environmental-politics-one-model-to-rule-them-all/.

model of long-term environmental cost incentives. In this model, higher numbers correspond to greater environmental damage. The representation of degrees in the center of each square represents theoretical global outcomes based on actions of each country, according to the global thresholds established in the Paris Agreement: the best-case scenario is to cap warming at 1.5°C below pre-industrial levels, while over 3°C of warming will result in the worst outcome.[29]

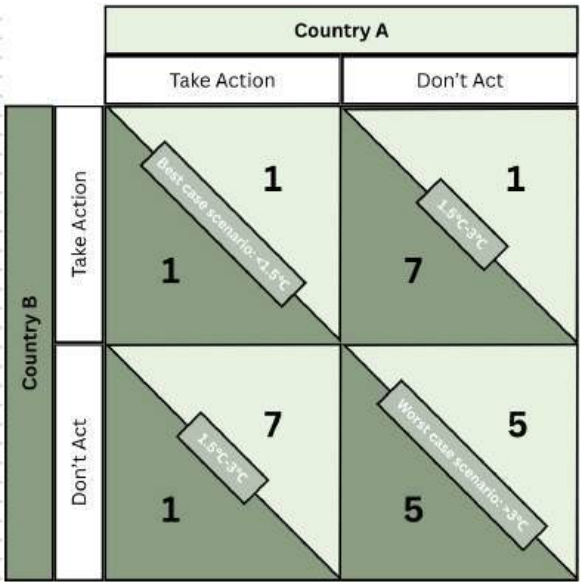


Figure 2

Mutual cooperation produces the lowest amount of environmental harm, while mutual defection produces the worst case of environmental consequences. When one country takes action while the other does not, the country that acts will theoretically face greater climate impacts, since it would have fewer resources for adaptation. Unlike the short-term model, cooperation is clearly the optimal collective outcome. However, the incentive to defect remains, as each country can avoid the costs of action while still benefiting from others' efforts.

This conflict of interest between short-term costs and long-term collective benefits is evident in the context of climate change.[30] Personal gains, in this case, consist of short-term convenience and comfort, whereas collective benefits are climate change mitigation and environmental stability. Reducing emissions incurs immediate, measurable costs, while the benefits of these actions are immeasurable, broadly distributed, and gradually visible over longer periods of time. Each country benefits from the mitigation efforts of others while minimizing the importance of its own contributions, making defection the rational choice for each nation. This logic is crucial to understanding climate inaction; cooperation is not irrational because of a lack of care, but because it is dependent on others' actions in ways that make it unreliable. Defection,

[29] United Nations Climate Change, "The Paris Agreement."
[30] CORE Econ, *The Economy*.

therefore, becomes the safer and more predictable strategy, even though countries would be better off if they all cooperated.

The structure of the Paris Agreement reflects an attempt to solve a global prisoner's dilemma using communication and peer pressure. While all countries collectively benefit from reducing emissions, short-term economic costs incentivize countries to minimize their own costs while still benefiting from the mitigation effects of others. In this way, the Agreement seems to be built on the hope that voluntary cooperation can overcome these misaligned incentives. Nonetheless, the underlying structure of the problem remains unchanged: cooperation is costly, benefits are shared globally, and no single country can guarantee others' cooperation. Therefore, the same dynamics that define the prisoner's dilemma, especially incentives to defect, continue to shape outcomes under the Paris Agreement.

How Countries Respond to this Incentive Structure: Temporal Discounting, Uncertainty, and Externalities

While the prisoner's dilemma explains the structural incentives that drive climate inaction, it does not fully account for how countries interpret and respond to these incentives. In reality, decisions are not purely rational but are shaped by factors that influence how governments perceive time, impact, and risk. These factors help to explain why countries often fail to act even when they understand the consequences of climate change and acknowledge the benefits of collective action.

Temporal discounting, defined as the tendency to prioritize short-term gains over distant consequences, plays a central role in climate inaction.[31] In deciding climate policy and impact, governments must weigh the immediate costs of mitigation against long-term benefits. However, the benefits of these actions are delayed and uncertain, making them seemingly less compelling than short-term costs. The benefits of mitigating the effects of climate change are long-term and less openly visible, lacking the immediacy and visibility that often drive decision-making. As a result, governments visibly see a weak connection between action and outcome, making the justification of short-term economic sacrifices for distant benefits more difficult.[32] At the national level, this tendency is closely tied to present bias, which refers to immediate costs being weighed more heavily than future gains.[33] In climate policy, this results in underinvestments in mitigation efforts, as governments often delay action to maintain economic growth.[34]

In addition to the discounting of long-term benefits, uncertainty further complicates climate decision-making. The uncertainty evident in climate change is not simply a matter of

[31] Jiyoung Kim, Yoonjae Nam, and Jihyun Yeo, "The Effects of Temporal Discounting on Perceived Seriousness of Environmental Behavior: Exploring the Moderator Role of Consumer Attitudes Regarding Green Purchasing," *Sustainability* 13, no. 13 (2021): 7130, https://www.mdpi.com/2071-1050/13/13/7130.

[32] Kim, Nam, and Yeo, "The Effects of Temporal Discounting on Perceived Seriousness of Environmental Behavior," 7130.

[33] Ismayil Penjaliyev, "Present Bias vs. Future Gains: Temporal Discounting in Climate Decision-Making," *ResearchGate*, May, 2025, https://www.researchgate.net/publication/392080520_Present_Bias_vs_Future_Gains_Temporal_Discounting_in_Climate_Decision-Making

[34] Penjaliyev, "Present Bias vs. Future Gains."

risk, but of a deeper ambiguity. Risk refers to situations where the possibilities are known, whereas uncertainty refers to situations where the possibilities are unclear.[35] In climate change, not only are future outcomes uncertain, but so is the likelihood of climate-related catastrophes. This makes the consequences of action versus inaction more difficult to evaluate, driving cooperation further and further away from seeming rational. Consequently, decision-making becomes less about optimization and more about perception and assumptions. This weakens rationality and increases the likelihood of defaulting to inaction.

In addition to uncertainty regarding environmental outcomes, there is also uncertainty about the actions of others.[36] Even if the consequences could be fully understood, uncertainty about others' cooperation remains a huge barrier. Players in the prisoner's dilemma are unable to accurately predict the actions of others, which is also the case in the context of climate change. Countries form expectations based on limited information, past inaction, weak enforcement, and inconsistency.[37] This creates low trust in cooperation; when countries expect others to defect, cooperation becomes risky rather than rational, essentially producing a self-reinforcing cycle of inaction.

This dynamic is further reinforced by the long-term and repeating nature of climate change. Unlike a one-time interaction or decision, the climate changes over time, resembling a repeated, dynamic game. Environmental outcomes are shaped by cumulative emissions, and the positive effects of action are often delayed and diffused.[38] This creates a system in which the visible feedback that would normally encourage behavioral change is weakened. Climate change can be understood as a repeated strategic interaction, where outcomes depend on how players respond to each other over time. In theory, repeated interactions could encourage cooperation through credible threats.[39] However, this would require consistent monitoring, clear consequences for defection, and trust in enforcement mechanisms–conditions that are often weak in climate governance.

The economic structure of climate change further reinforces this incentive system through the presence of externalities. In economics, an externality is a cost or benefit arising from an activity that impacts a third party who is not directly involved.[40] Climate change is driven by negative externalities, where environmental costs of emissions are not directly reflected in market prices.[41] Since no formal system assigns prices to environmental harm,

---

[35] Jose A. Rodreigues-Neto, "Climate Change Policy: A Theorist's Plea to Take Heed of Game Theory and Ambiguity Aversion," *Agenda: A Journal of Policy Analysis and Reform* 15, no. 2 (2008): 85-90, https://www.jstor.org/stable/43199521?seq=1.

[36] Rodrigues-Neto, "Climate Change Policy."

[37] Rodrigues-Neto, "Climate Change Policy."

[38] CORE Econ, *The Economy*.

[39] Jobst Heitzig, Kai Lessmann, and Yong Zou, Self-Enforcing Strategies to Deter Free-Riding in the Climate Change Mitigation Game and Other Repeated Public Good Games," *Proceedings of the National Academy of Sciences* 108, no. 38 (2011): https://www.pnas.org/doi/10.1073/pnas.1106265108.

[40] Zili Yang, "Climate Change and Externality," *Climate Change Economics* 11, no. 4 (2020): 1-9, https://www.jstor.org/stable/27130726?seq=1.

[41] CORE Econ, *The Economy*.

individuals and firms are not required to fully account for the environmental costs of their actions. Consequently, high carbon emissions remain economically rational to nations, even though they produce harmful consequences collectively. Market incentives, therefore, favor short-term economic gains over long-term environmental stability, creating structural conditions that continuously incentivize unsustainable behavior. Fossil fuels, for example, appear cheap compared to cleaner alternatives because their environmental costs are externalized. When a company burns fossil fuels for its own benefit, the environmental damage caused by emissions affects society as a whole. Since the damage is not paid directly by the company, it is considered an external cost. Market prices reflect production costs; however, environmental damage is rarely included. Since businesses do not pay the environmental costs of these emissions, the market ultimately reinforces this behavior, rewarding short-term efficiency while ignoring long-term collective harm.

Ultimately, the prisoner's dilemma provides a useful framework for understanding why climate cooperation is extremely difficult to achieve. It demonstrates how rational decision-making can produce outcomes that are collectively irrational. When combined with temporal discounting, uncertainty, and externalities, this structure becomes even more resistant to cooperation. As a result, climate inaction is not simply a failure of cooperation, but instead a rational response to a system in which short-term incentives, imperfect information, and delayed outcomes discourage meaningful action.

Competing Incentives and Unequal Payoffs: The European Union vs. India Case Study

The tension between short-term economic incentives and long-term environmental outcomes is reflected in real-world climate policy. As illustrated in Figures 1 and 2, countries are playing two simultaneous games: a short-term economic game, where defection lowers immediate costs, and a long-term environmental game, where cooperation produces the best collective outcome.[42] The structure of this dilemma remains the same; however, the payoffs differ across different countries due to economic conditions. This interaction becomes especially visible when comparing developed and developing countries, where economic positions establish discrepancies in their trade-offs.

A clear example of this can be seen by comparing the European Union to India. While both are participants in the Paris Agreement and have formally committed to taking action, their priorities vary significantly. As a developed, high-income region, the EU has had greater financial capacity to prioritize long-term environmental goals, already having industrialized and built much of its infrastructure. India, as a developing country, faces stronger pressures to prioritize industrial growth, which makes the immediate costs of reduction more significant. Therefore, the same climate decision carries significantly different consequences for each nation.

Figure 3 below is a modified version of Figure 1, illustrating short-term costs of climate mitigation for the EU and India. In the short term, climate mitigation requires immediate economic costs, including transitioning energy systems, infrastructure, and potentially slowing

[42] Magli and Manfredi, "Coordination Games vs Prisoner's Dilemma in Sustainability Games."

economic growth. Not taking action allows countries to maintain economic growth, while taking action takes resources away from development. Greater numbers in this model represent a greater cost burden, as similar to Figure 1.

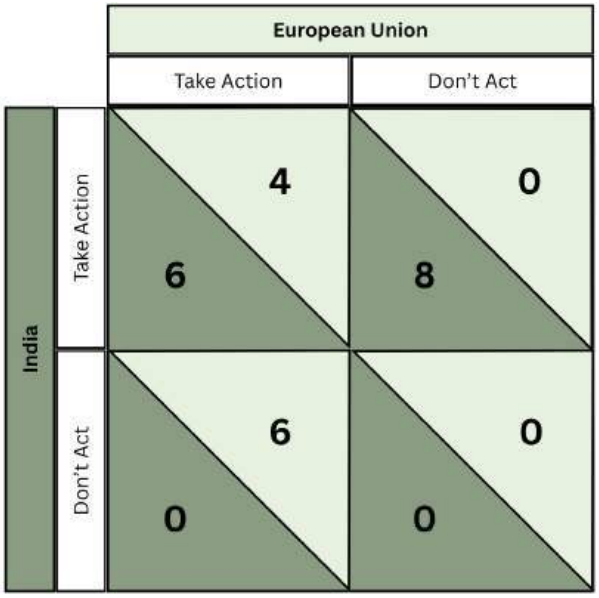


Figure 3

The differences in numbers within this model reflect the differences in economic capacity for each country. As a developing country, India continues to rely heavily on fossil fuels to sustain economic growth. Fossil fuels account for around 75% of India's total energy demand, with coal alone providing the majority of electricity.[43] This reliance reflects the role of low-cost energy sources in supporting industrialization. Allocating resources towards emissions reductions represents a significant opportunity cost for India, as these same resources could have otherwise been used for economic development.[44]

In contrast, the European Union is better positioned to bear immediate economic costs. The EU has already invested heavily in climate mitigation, implementing policies such as the Emissions Trading System, the world's first and largest "cap and trade" carbon market.[45] The EU's most recent NDC, updated in 2025, emphasizes goals of "achieving a net reduction of 55% in greenhouse gas (GHG) emissions by 2030 and introduces an indicative contribution of 66.25% to 72.5% for 2035 on the path towards carbon neutrality by 2050."[46] The EU is currently on track

[43] Climate Action Tracker, "Breaking the Cycle of Energy Shocks: India's Renewable Opportunity," April 15, 2026, https://climateactiontracker.org/blog/breaking-the-cycle-of-energy-shocks-indias-renewable-opportunity/.

[44] Josie Garthwaite, "The Opportunity Costs of Carbon Capture," Stanford Doerr School of Sustainability, February 13, 2025, https://sustainability.stanford.edu/news/opportunity-costs-carbon-capture.

[45] European Commission, "About the EU Emissions Trading System," *Climate Action*, accessed April 22, 2026, https://climate.ec.europa.eu/eu-action/carbon-markets/about-eu-ets_en.

[46] Council of the European Union, "Paris Agreement: The EU Submits its Updated NDC with an Indicative Target for 2035 to the UN Ahead of COP30," November 5, 2025,

to meet this target, demonstrating how its policy and existing infrastructure make its mitigation goals achievable.[47] Since most of its infrastructure is already developed, diverting resources towards climate action does not come at the same expense of economic growth as it does for India. For India, the cost of cooperation is higher, making defection the rational strategy. For the EU, the cost is lower, making cooperation more feasible. However, defection remains individually rational for both countries, as it avoids immediate economic losses.

When the same decision of action versus inaction is evaluated over a long-term environmental lens, the incentives are reversed. Figure 4 below is a modified version of Figure 2, illustrating the long-term environmental impacts of action versus inaction, rather than immediate economic cost. The greater the number, the worse the environmental damage. Here, defection produces the worst collective outcome, as continued emissions lead to increased negative climate impacts, while cooperation minimizes this damage.

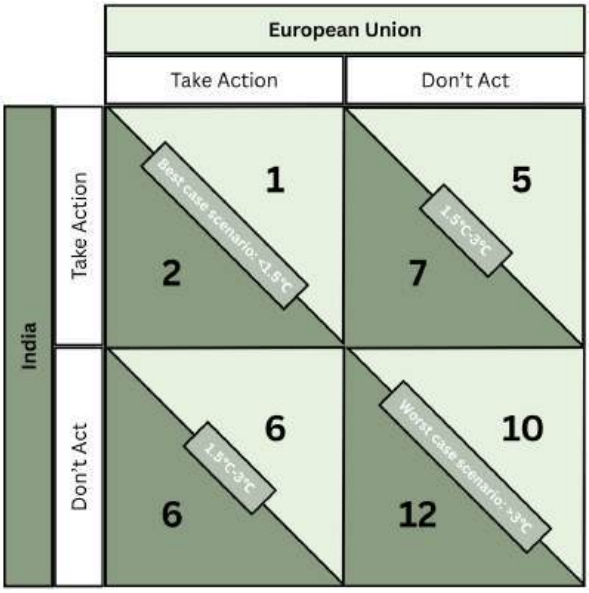


Figure 4

Within this long-term model, the distribution of impacts remains unequal but remains collectively worsened by defection. India, despite contributing less historically to global emissions, is the seventh most vulnerable country to the impacts of climate change.[48] It faces greater risks from heatwaves, flooding, and agricultural disruption. Despite this vulnerability, India continues to prioritize development, as immediate economic growth creates stronger short-term incentives that often outweigh uncertain long-term environmental benefits. The

https://www.consilium.europa.eu/en/press/press-releases/2025/11/05/paris-agreement-the-eu-submits-its-updated-ndc-with-an-indicative-target-for-2035-to-the-un-ahead-of-cop30/.

[47] European Commission, "EU is Progressing Towards its 2030 Climate and Energy Goals while Tackling High Energy Prices, State of the Energy Union Report 2025 Shows," November 5, 2025, https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2586.

[48] World Economic Forum, "India is Showing Developing Economies How to Move from Climate Ambition to Action," January 14, 2025, https://www.weforum.org/stories/2025/01/india-climate-opportunity-developing-economies/.

European Union, while also exposed to risks, has a greater economic capacity to mitigate them, which reduces the cost of climate action compared to developing economies.

Although cooperation is clearly optimal in the long-term model, countries do not make decisions solely based on this framework. Instead, they simultaneously make decisions under both long-term and short-term models, where immediate economic incentives often tend to dominate decision-making. Together, these models demonstrate why climate cooperation is extremely difficult to achieve. In the long-term, cooperation is rational, mitigating climate change for all. But, defection is rational in the short-term, especially for developing countries such as India. This comparison ultimately suggests that the prisoner's dilemma is not a single, static game in the context of climate change, but instead a layered one of competing incentives over time.

Reframing Climate Action Incentives

Overall, the value-action gap demonstrates that widespread awareness does not guarantee meaningful action. Climate change inaction is not primarily due to a lack of awareness; it instead reflects a deeper misalignment between how short-term incentives are structured and how long-term decisions are made. The prisoner's dilemma explains why rational players are incentivized to defect in the short-term rather than cooperate, reinforcing prioritizing immediate benefits and weakening incentives for cooperation.

Understanding climate change through the combination of these factors helps to explain why even highly concerned players fail to act. As long as short-term incentives outweigh long-term benefits and governments remain uncertain about the impacts of their actions, the value-action gap will likely persist. Addressing climate change, therefore, requires significant modifications regarding incentives, uncertainty, and accountability. Without such changes, the same dynamics that have historically limited cooperation could thereby lead to an outcome that adversely impacts life on Earth.

Works Cited

European Commission. “About the EU ETS - Climate Action.” Accessed April 23, 2026. https://climate.ec.europa.eu/eu-action/carbon-markets/about-eu-ets_en.

Alessio, Carrozzo Magli, and Piero Manfredi. “Coordination Games vs Prisoner’s Dilemma in Sustainability Games: A Critique of Recent Contributions and a Discussion of Policy Implications.” Ecological Economics 192 (February 2022): 107268. https://doi.org/10.1016/j.ecolecon.2021.107268.

Cohen, Li. “Global Climate Disaster Inevitable If Emissions Aren’t Drastically Reduced by 2035, U.N. Warns.” CBS News, October 25, 2024. https://www.cbsnews.com/news/climate-change-un-report-disaster-emissions-future/.

Council of the European Union. “Paris Agreement: The EU Submits Its Updated NDC with an Indicative Target for 2035 to the UN Ahead of COP30.” November 5, 2025. https://www.consilium.europa.eu/en/press/press-releases/2025/11/05/paris-agreement-the-eu-submits-its-updated-ndc-with-an-indicative-target-for-2035-to-the-un-ahead-of-cop3

Climate Action Tracker. “Countries.” Accessed April 23, 2026. https://climateactiontracker.org/countries/.

Das, Nandini. “Breaking the Cycle of Energy Shocks: India’s Renewable Opportunity.” Climate Action Tracker. March 26, 2026. https://climateactiontracker.org/blog/breaking-the-cycle-of-energy-shocks-indias-renewable-opportunity/.

Dwyer, Augusta. “6 Things That Have Happened since the Paris Agreement.” Think Landscape (Global Landscapes Forum). December 9, 2025. https://thinklandscape.globallandscapesforum.org/100713/10-years-paris-agreement/.

UN Environment Programme. Emissions Gap Report 2025. Nairobi: United Nations Development Programme. November 4, 2025. https://www.unep.org/resources/emissions-gap-report-2025.

European Commission. “The EU Is Progressing towards Its 2030 Climate and Energy Goals While Tackling High Energy Prices.” November 5, 2025. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2586.

Fagan, Moira, Kirsten Lesage, William Miner, and Andrew Prozorovsky. “People in Middle-Income Countries Say Climate Change Is Affecting Their Community.” Pew Research Center, November 5, 2025. https://www.pewresearch.org/global/2025/11/05/people-in-middle-income-countries-say-climate-change-is-affecting-their-community/.

Farias, Ana Rita, Seray Coruk, and Cláudia Simão. “The Effects of Temporal Discounting on Perceived Seriousness of Environmental Behavior: Exploring the Moderator Role of Consumer Attitudes Regarding Green Purchasing.” Sustainability 13, no. 13 (2021): 7130. https://doi.org/10.3390/su13137130.

Garthwaite, Josie. “The Opportunity Costs of Carbon Capture.” Stanford Doerr School of Sustainability,” February 13, 2025. https://sustainability.stanford.edu/news/opportunity-costs-carbon-capture.

Grunwald, Guido, Ali Kara, and John E. Spillan. “Involvement Matters: Navigating the Value–Action Gap in Business Students’ Sustainability Transformation Expectations—a Cross-Country Kano Study.” Sustainability Science (2025). https://link.springer.com/article/10.1007/s11625-025-01681-z

Heitzig, Jobst, Kai Lessmann, and Yong Zou. “Self-Enforcing Strategies to Deter Free-Riding in the Climate Change Mitigation Game and Other Repeated Public Good Games.” Proceedings of the National Academy of Sciences 108, no. 38 (2011): 15739-15744 https://www.pnas.org/doi/10.1073/pnas.1106265108.

Hickel, Jason. Less is More: How Degrowth Will Save The World. London: Windmill Books, 2021.

Penjaliyev, Ismayil. “Present Bias vs. Future Gains: Temporal Discounting in Climate Decision-Making.” May 2025. https://www.researchgate.net/publication/392080520_Present_Bias_vs_Future_Gains_Temporal_Discounting_in_Climate_Decision-Making

Rodrigues-Neto, José A. “Climate Change Policy: A Theorist’s Plea to Take Heed of Game Theory and Ambiguity Aversion.” Agenda: A Journal of Policy Analysis and Reform 15, no. 2 (2008): 85–90. https://www.researchgate.net/publication/266294083_Climate_Change_Policy_A_Theorist's_Plea_to_Take_Heed_of_Game_Theory_and_Ambiguity_Aversion

Shah, Anish. “India Is Showing Developing Economies How to Take Action on the Climate.” World Economic Forum, January 14, 2025. https://www.weforum.org/stories/2025/01/india-climate-opportunity-developing-economies/.

Shatz, Itamar. “The Value-Action Gap: Why People Don’t Act in Accordance with Their Beliefs.” Effectiviology. Accessed April 23, 2026. https://effectiviology.com/value-action-gap/.

Sterescu, Andreï-Bogdan. “The Prisoner’s Dilemma in Environmental Politics: One Model to Rule Them All? – E-International Relations.” E-International Relations, June 14, 2018. https://www.e-ir.info/2018/06/14/the-prisoners-dilemma-in-environmental-politics-one-model-to-rule-them-all/.

CORE Econ. The Economy 1.0. Oxford: Oxford University Press, 2017. https://books.core-econ.org/the-economy-v1/en/.

Tso, Kathryn. “How Are Countries Held Accountable under the Paris Agreement? | MIT Climate Portal.” MIT Climate Portal, August 8, 2025. https://climate.mit.edu/ask-mit/how-are-countries-held-accountable-under-paris-agreement.

United Nations. “The Paris Agreement.” United Nations Climate Change. Accessed April 23, 2026. https://www.un.org/en/climatechange/paris-agreement.
United Nations Climate Change. “The Paris Agreement.” Accessed April 23, 2026. https://unfccc.int/process-and-meetings/the-paris-agreement.
United Nations Development Programme. “The World’s Largest Survey on Climate Change Is out – Here’s What the Results Show.” UNDP Climate Promise, June 27, 2024. https://climatepromise.undp.org/news-and-stories/worlds-largest-survey-climate-change-out-heres-what-results-show.
Yeo, Sophie. “Explainer: The ‘Ratchet Mechanism’ within the Paris Climate Deal.” Carbon Brief, December 3, 2015. https://www.carbonbrief.org/explainer-the-ratchet-mechanism-within-the-paris-climate-deal/.

Assessing the Role of In Vitro Fertilization in South Korea's Demographic Crisis
By Gloria Kim

Introduction

2023 signaled a precarious year for South Korea's future. Then, the country just surpassed its own record for the world's lowest fertility rate (FR) from the previous year, reaching the lowest-ever average birth rate recorded in history, 0.72 children per woman (Choi). Since 2020, the country has seen a sharp decline in fertility, with a natural population contraction over five consecutive years. However, as the Seoul Metropolitan Government increased subsidies for fertility treatments by 61.8%, for the first time in a decade, Korea's FR rebounded to 0.75 in 2024 and later continued to reach 0.79 (Choi). This paper analyzes the role of In Vitro Fertilization (IVF) in revolutionizing fertility outcomes in South Korea, considering the root causes of South Korean infertility and prevalent maternal health risks.

The paper will be guided by the following research question: Is South Korea's growing reliance on IVF medically justified for subfertile women in terms of reproductive success, maternal health risks, and sustainability? Only after evaluating the integrity of IVF's outcomes on fertility and maternal health can its long-term performance be predicted.

Causes of South Korean Infertility

In South Korea, where the nuptiality and fertility rates have plummeted for the past few decades, the postponement of marriage has resulted in the delay of childbearing. In 1994, the average age at first birth was 26.4; in 2024, the number rose to 33.1 (Kim). Consequently, the leading cause of infertility in South Korea can be identified as advanced maternal age, where any maternal age above 35 poses a statistically significant risk (Lee et al.). In spite of this, it is worth noting that from 2022 to 2024, male infertility cases increased by 37% from 79,176 to 108,358, while female cases only rose by 29% from 149,442 to 192,043 (Hong), demonstrating the limitations of pinpointing female infertility alone. Although male infertility currently accounts for 30-50% of all infertility cases in South Korea ("Male Fertility"), it may take on a larger role in contributing to the nation's low birth rate in the future.

At present, advanced maternal age remains the predominant biological explanation for infertility, which results from diminished ovarian reserve, an inevitable physiological transition in all women approaching menopause. Ovarian reserve is defined as the total count of healthy oocytes, or immature eggs capable of fertilization, remaining in a woman's ovaries (Ulrich and Marsh). All females are born with a finite oocyte reserve and have their peak count in the mother's uterus. From this stage onward, females experience oocyte atresia, the gradual degeneration of immature oocytes within the ovary throughout the lifespan until menopause, signaling complete depletion of the oocytes. As a woman ages, her ovarian reserve declines in both number and quality, leading to aneuploidy, a condition in which a gamete has an irregular number of chromosomes due to errors in cell division (Mikwar et al.). Furthermore, aneuploidy is a major cause of failed implantation of the embryo, miscarriage, as well as congenital birth

defects and developmental disabilities such as Down syndrome ("Aneuploidy"). Regardless of natural or artificial conception, the risk for aneuploidy drastically increases for women after 35.

Despite the issue of male infertility contributing to the low fertility rate, advanced maternal age, combined with increased accessibility to modern technology, has driven an immense surge in IVF usage. In fact, ovarian aging has developed into reliance on IVF, where 63.5% of subfertile women in South Korea who utilize Artificial Reproductive Technologies (ARTs) to conceive are above 35 years old. In 2022, the distribution of women's ages undergoing fertility treatment showed that the largest cohort was between 35-39 at 37.4%, followed by women between 40-44 (26.1%) (Lee et al.). In addition, the proportion of babies conceived through IVF has surged by 83.2% since 2020, and in 2024, 20.2% of babies were born from successful fertility treatment ("Seoul Expands"), showing the growing use of ARTs in a subfertile population battling old age (Yoon). In sum, the latter has revolutionized fertility outcomes in South Korea due to its unique ability to partially circumvent the biological limitations of ovarian ageing, thus enabling conception even among women at advanced ages, unlike other ARTs that are strictly constrained by maternal age.

Maternal Psychological Implications

While IVF offers a biomedical solution to age-related infertility, over-reliance has introduced significant psychological consequences for patients. In a study observing the extent of depression amongst South Korean women using fertility treatment, 84.2% showed certain depressive disorders, assessed using the Center for Epidemiological Studies Depression (CES-D) scale (Kim et al.). However, the direct psychological impact of treatment remains difficult to quantify consistently, as another study conducted by Professor Park at Kyungpook National University found that only 12.9% of IVF patients experienced clinically substantial anxiety with a Hospital Anxiety and Depression Scale (HADS) score greater than 11 out of 21, while 32.3% experienced depression (>11). (Park) The CES-D scale and HADS are both self-report questionnaires designed to clinically evaluate anxiety and depression; inconsistent results evidence the need for standardization in data collection methods.

Nonetheless, IVF's frequent clinical interventions are psychologically taxing for the patients, which include invasive procedures such as, but not limited to: hormonal injections, blood retrievals, and daily ultrasounds. With greater disruptions in their daily functioning, women experience a decreased quality of life than their male counterparts, surrounding treatment-related worries and fear of failure that are exacerbated by self-blame and maladaptive coping strategies (Grammenou et al.). When an individual experiences infertility, marital intimacy, communication, and interpersonal dynamics in a couple are severely impacted (Holland et al.). According to Park (2025), a lack of intimacy between couples results in psychological maladjustment and distress at the individual level. Indeed, inadequate conjugal support may establish a perpetuating cycle of infertility by permanently discouraging women from continuing fertility treatment.

In a quasi-experimental study, results revealed that a psychological intervention program for women undergoing ART considerably alleviated depression and fatigue, through targeting marital intimacy and sexual fulfillment (Kim et al.). Similarly, a comprehensive review by Dr. Arzu Yurci at Istanbul Arel University in Turkey (2024) highlighted the positive correlation between mental intervention programs, such as counseling, stress coping programs, and cognitive-behavioral therapy, and higher IVF conception rates, indicated by higher conception and lower miscarriage rates in patients compared to control groups. Therefore, a well-integrated and accessible mental support system in South Korean fertility care would significantly decrease a major component of maternal health risks associated with IVF, enabling South Korea to continue to utilize IVF to sustainably boost its fertility rate.

Conclusion

Is South Korea's growing reliance on IVF medically justified for subfertile women in terms of reproductive success, maternal health risks, and sustainability? This paper identifies advanced maternal age as the root cause behind South Korea's low fertility rate and the increasing reliance of couples on IVF to boost their chances of conception at an advanced age. Although IVF has yielded a large proportion of national live births, it does not address the underlying causes of advanced maternal age and declining ovarian reserve. Treatment does not reverse ovarian maturation, reduce biological age, nor decrease the risk of aneuploidy; moreover, IVF treatment severely hinders the quality of life of patients, and can cause psychological implications by disrupting marital intimacy and sexual function in couples. However, when accompanied with psychological interventions, such as counseling and group therapy sessions, IVF can significantly increase the chances of successful implantation, resulting in higher live birth rates. At a national scale, these individual consequences may significantly alter broader demographic patterns.

As IVF continues to dominate the lives of South Korean couples, its bioethical justification must not only depend on its numerical success as a national fertility strategy, but rather hinge on the integration of psychological support programs for patients to ensure that reproductive gains do not come at maternal health costs.

Works Cited

"Aneuploidy." Cleveland Clinic, 25 Aug. 2022, my.clevelandclinic.org/health/diseases/24060-aneuploidy. Accessed 1 Mar. 2026.

Choi, Seul-Ki. Korean Social Trends. Ministry of Data and Statistics Data and Research Institute, 1 Jan. 2023, mods.go.kr/board.es?mid=b10104000000&bid=12046&act=view&list_no=432665. Accessed 1 Mar. 2026.

Grammenou, Maria, et al. "The Psychological Impact of in Vitro Fertilization (IVF): A Gender Systematic Review." Healthcare, vol. 14, no. 3, 2 Feb. 2026, p. 375, https://doi.org/10.3390/healthcare14030375.

Holland, Karen J., et al. "Spiritual Intimacy, Marital Intimacy, and Physical/psychological Well-being: Spiritual Meaning as a Mediator." Psychology of Religion and Spirituality, vol. 8, no. 3, 2016, pp. 218-27, https://doi.org/10.1037/rel0000062. Accessed 1 Mar. 2026.

Kim, Miok, et al. "Effects of Psychological Intervention for Korean Infertile Women under in Vitro Fertilization on Infertility Stress, Depression, Intimacy, Sexual Satisfaction and Fatigue." Archives of Psychiatric Nursing, vol. 34, no. 4, Aug. 2020, pp. 211-17, https://doi.org/10.1016/j.apnu.2020.05.001.

Lee, Dayong, et al. "Assisted Reproductive Technology Trends in Korea: Annual Report for 2022." Clinical and Experimental Reproductive Medicine, vol. 52, no. 4, 1 Dec. 2025, pp. 299-309, https://doi.org/10.5653/cerm.2025.08872.

"Male Fertility Clinic." Cha Fertility, Cha University Global IVF Group, seoul.chamc.co.kr/en/clinic/male.aspx#:~:text=When%20the%20cause%20of%20infertility,fertility%20in%20male%20infertility%20patients. Accessed 1 Mar. 2026.

"Male infertility cases in Korea top 100,000 amid overall infertility surge." Chosun Biz En, 19 Oct. 2025, biz.chosun.com/en/en-science/2025/10/19/J4MYS6ND3FCNJHI3ONSR4OA23M/. Accessed 1 Mar. 2026.

Mikwar, Myy, et al. "Mechanisms of Oocyte Aneuploidy Associated with Advanced Maternal Age." Mutation Research/Reviews in Mutation Research, vol. 785, July 2020, p. 108320, https://doi.org/10.1016/j.mrrev.2020.108320.

Park, Joon Cheol. "Assessment of Anxiety, Depression, and Sexual Dysfunction in Women Undergoing Fertility Treatment." Clinical and Experimental Reproductive Medicine, vol. 52, no. 1, 1 Mar. 2025, pp. 38-43, https://doi.org/10.5653/cerm.2024.07465.

“Seoul Expands Support for Infertility Treatments: 1 in 5 Births in Seoul Results from Infertility Procedure.” Seoul Metropolitan City, Seoul Metropolitan Government, 4 Nov. 2024, english.seoul.go.kr/seoul-expands-support-for-infertility-treatments-1-in-5-births-in-seoul-results-from-infertility-procedure/. Accessed 1 Mar. 2026.

"South Korean Marriage Age Delayed 6-7 Years vs. Parents' Generation." The Chosun Daily, 19 Feb. 2026,

www.chosun.com/english/market-money-en/2026/02/12/3VPIXR7RF5C2XJQ4OXSY5AQZKU/. Accessed 1 Mar. 2026.

Ulrich, Nicole D., and Erica E. Marsh. "Ovarian Reserve Testing: A Review of the Options, Their Applications, and Their Limitations." Clinical Obstetrics & Gynecology, vol. 62, no. 2, June 2019, pp. 228-37, https://doi.org/10.1097/grf.0000000000000445. Accessed 1 Mar. 2026.

Yoon, Min-sik. "In vitro fertilization nearly doubles in 5 years." The Korea Herald, 17 july 2025, www.koreaherald.com/article/10533988#:~:text=July%2017%2C%202025%20%2D%2014. Accessed 1 Mar. 2026.

Yurci, Arzu. "The Impact of Psychological Factors on in Vitro Fertilization (IVF) Outcomes: A Comprehensive Review." Journal of Comprehensive Surgery, vol. 2, no. 2, 29 May 2024, pp. 35-38, https://doi.org/10.51271/jocs-0031.

Extreme Heat Is a Public Health Emergency in Massachusetts By Deetya Prasad

Massachusetts is often viewed as a cold-weather state. Long winters and relatively short summers can make extreme heat seem like a distant concern. That perception is becoming increasingly inaccurate. Rising temperatures and more frequent heat waves have made extreme heat a growing public health threat across the Commonwealth. While Massachusetts has spent decades developing policies and infrastructure to protect residents from severe winter weather, it has not adapted as quickly to the dangers posed by extreme heat. As temperatures continue to rise, treating heat as an occasional weather event rather than a public health emergency places residents at increasing risk.

Recent data demonstrate a clear warming trend across the state. According to a 2025 analysis published by CommonWealth Beacon, average summer temperatures in Massachusetts have increased by approximately 2.75°F since 1970 (Borkhetaria). Scientists expect extreme heat events to become more common as climate change continues. Although heat is often associated with states such as Arizona or Texas, Massachusetts is already experiencing the effects of rising temperatures. Extreme heat is no longer an occasional anomaly. It is becoming a regular feature of summer in the Commonwealth.

These changes have important consequences for public health. The human body relies on sweating and increased blood flow to regulate temperature. When temperatures become too high, these mechanisms can fail. Heat exhaustion can develop after prolonged exposure to heat and dehydration. If untreated, it can progress to heat stroke, which is a life-threatening medical emergency. The Centers for Disease Control and Prevention identifies heat as the leading cause of weather-related deaths in the United States ("Heat & Health Tracker"). Heat also contributes to dehydration, cardiovascular complications, and worsening symptoms among individuals with chronic medical conditions. As temperatures rise, the health impacts of heat place increasing pressure on hospitals, emergency departments, and public health systems.

Massachusetts residents are already experiencing these consequences. Data from the Massachusetts Environmental Public Health Tracking Program show that heat-related illness continues to result in hospitalizations across the state ("Heat Stress Hospitalization"). These figures likely underestimate the true burden because they capture only individuals who receive medical care. Many people experience heat-related symptoms without seeking treatment. Public health officials also warn that heat-related illness is often underreported because heat can worsen existing medical conditions rather than appear as the primary cause of illness or death ("Extreme Heat").

The burden of extreme heat is not shared equally. Older adults, young children, pregnant individuals, outdoor workers, and people with chronic medical conditions face elevated risks during heat waves. Massachusetts has a large and growing older population, which increases the number of residents vulnerable to heat-related illness. Low-income households also face greater risks because many residents cannot afford adequate cooling during prolonged heat events. Neighborhood conditions can further intensify exposure. Areas with limited tree cover and large

amounts of pavement often experience higher temperatures than surrounding communities, a phenomenon known as the urban heat island effect. The City of Boston has identified significant temperature differences across neighborhoods due to variations in tree canopy and land use ("Heat Resilience Solutions for Boston").

Chelsea illustrates how these disparities can affect communities. Research examining residents of Chelsea and East Boston found that many people face barriers to staying cool during extreme heat events, including housing conditions, limited access to cooling resources, and concerns about energy costs (McIntyre et al.). These challenges increase exposure to dangerous temperatures and can worsen existing health inequities. As a result, extreme heat often has its greatest impact on populations that already experience disproportionate health burdens.

Despite these growing risks, Massachusetts remains better prepared for extreme cold than for extreme heat. State policies have long recognized the dangers of cold exposure and require landlords to maintain minimum indoor temperatures during the winter months. Comparable protections for dangerously high indoor temperatures are far less common. This gap reflects an outdated understanding of climate risk. Many housing regulations were developed when extreme heat was less frequent in New England. Today, however, summers are becoming hotter while many buildings remain poorly equipped to handle prolonged heat waves. Residents may be protected from unsafe indoor temperatures in January while facing dangerous indoor conditions in July.

Existing responses also have important limitations. Massachusetts issues heat advisories and provides guidance during extreme heat events ("Extreme Heat"). Many communities also operate cooling centers during periods of dangerous temperatures. These measures provide valuable assistance, but they do not reach everyone who needs help. Older adults may face transportation barriers, and outdoor workers often cannot avoid prolonged exposure. Heat emergencies, therefore, require more than short-term responses. They require policies that reduce vulnerability before temperatures become dangerous.

Massachusetts should begin treating extreme heat as a public health emergency and plan for it accordingly. State agencies should strengthen heat action plans and expand outreach to high-risk populations. Municipal governments should invest in tree planting, shade infrastructure, and neighborhood cooling strategies in areas that experience the highest temperatures. Building codes should encourage heat-resilient construction and energy-efficient cooling systems, particularly in older housing stock. Policymakers should also consider stronger protections for renters and workers during extreme heat events. These measures would reduce exposure, improve health outcomes, and help communities adapt to a changing climate.

Extreme heat is no longer a future concern for Massachusetts. It is already affecting communities across the state and placing vulnerable residents at risk. The evidence is clear, and effective solutions are available. Recognizing extreme heat as a public health emergency is an essential step toward protecting Massachusetts residents in a warming climate.

Works Cited

Borkhetaria, Bhaamati. "Average Summer Temps in Massachusetts Have Risen by 2.75 Degrees Fahrenheit in Past 55 Years, According to New Data." CommonWealth Beacon, 30 May 2025, https://commonwealthbeacon.org/environment/average-summer-temps-in-massachusetts-have-risen-by-2-75-degrees-fahrenheit-in-past-55-years-according-to-new-data/. Accessed 4 June 2026.

Centers for Disease Control and Prevention. "Heat & Health Tracker." CDC Environmental Public Health Tracking, https://ephtracking.cdc.gov/Applications/heatTracker/. Accessed 4 June 2026.

City of Boston. "Heat Resilience Solutions for Boston." City of Boston, 7 May 2025, https://www.boston.gov/departments/climate-resilience/heat-resilience-solutions-boston. Accessed 4 June 2026.

Massachusetts Department of Public Health. "Extreme Heat." Commonwealth of Massachusetts, https://www.mass.gov/extreme-heat. Accessed 4 June 2026.

Massachusetts Environmental Public Health Tracking Program. "Massachusetts Environmental Public Health Tracking." Commonwealth of Massachusetts, https://matracking.ehs.state.ma.us/Health-Data/index.html. Accessed 4 June 2026.

McIntyre, Alina M., et al. "Facilitators and Barriers for Keeping Cool in an Urban Heat Island: Perspectives from Residents of Chelsea and East Boston, Massachusetts." Environmental Justice, vol. 16, no. 1, 2023, https://www.liebertpub.com/doi/pdf/10.1089/env.2022.0019. Accessed 4 June 2026.

Energy Aware Terrain Mapping and Autonomous Recovery Landing of a UAV on A Moving Ground Dock Station By Almaz Sagnalin

Abstract

The low battery life is one of the main problems faced by UAVs, which limits the time for autonomous mapping and exploration missions. To overcome this problem, this paper offers a novel approach that utilizes an energy-aware terrain mapping combined with a predictive precision landing to the Mobile Ground Docking Station (MGDS) in motion. The flight plan of the drone is constantly updated depending on the current battery consumption, while the position of the drone is tracked via the Extended Kalman Filter (EKF) based on an AprilTag 3 marker without any camera delay during the descent. To prove the efficiency of our approach, we have conducted a total of 30 experiments in the real-life environment on the open field in the presence of the natural wind. The experiment has shown the efficiency of the system. During the experiment, the drone mapped an area of 14,000 square meters using 215 watt-hours of power. Furthermore, the landing was performed in 92% of cases with a mean error of just 4 centimeters and 16 seconds per landing on average. Finally, the mechanical locks installed on the drone were able to recover the drone in 89% of cases.

Introduction

A drone is able to map out the disaster scene, conduct infrastructure inspection or look for people who are missing without any assistance (Ullah & Ullah, 2026). But there is one thing that puts an end to the drone's mission in most cases, as it is exhausted due to running out of batteries (Nieuwoudt et al., 2023a). As the time passes, UAVs get better and are expected to show higher results at harder missions (Ullah & Ullah, 2026). Modern drones can make high quality images, 3D terrain maps and navigate through difficult obstacles. Nevertheless, limited battery capacity remains one of the biggest obstacles to long-timed autonomous operations (Kwak & Sung, 2018a). To address this challenge, researchers are exploring methods that allow UAVs to use energy more efficiently while also having automatic recovery and recharging at the docking station (Nieuwoudt et al., 2023b). This research focuses on combining energy aware terrain mapping with autonomous landing on a moving ground docking station to improve mission endurance and reliability. Autonomous drones have become impossible to replace in collecting geospatial and environmental data in situations with limited time, accessibility and fast changing operating conditions that present significant challenges (Ullah & Ullah, 2026). One of their mainest functions is terrain mapping, where aerial image from the drone is processed into detailed 3D maps of terrain. These abilities require continuous situational awareness, large area monitoring and making quick decisions. Recent upgrades in AI and computer vision have strengthened UAVs powers (Wiyagi et al., 2025). Modern systems can follow pre-written routes, avoid obstacles, determine the condition of the environment, and adapt to the situation with minimal human intention (Ghasemi et al., 2021). In parallel, researchers have developed automated docking and charging technologies that make the UAVs less dependent on humans

and give longer flight time (Nieuwoudt et al., 2023a). Combined, these developments get UAVs closer to fully autonomous operation without human supervision. However, despite these improvements the drone's limited energy remains a critical clamping that limits mission duration and range (Kwak & Sung, 2018b). The biggest limitation of most UAVs is their small energy capacity. Battery powered drones can stay airborne only for a very short time before being in need of a recovery or recharging (Nieuwoudt et al., 2023b). During autonomous mapping missions, running out of energy may force the early stop of data collection, leaving inspecting regions incomplete and reducing the overall results of the drone's work (Kwak & Sung, 2018a). In more critical situations, poor energy reserves can get a UAV to not reach a safe landing spot, causing the loss of collected data or even the loss of the aircraft itself. Although fixed docking stations can partially solve this problem. But they are often impractical in remote and continuously changing operating conditions where the drone's position and mission tasks change over time (Nieuwoudt et al., 2023a). Previous studies have investigated terrain mapping, autonomous navigation, and precision landing separately (Kwak & Sung, 2018a; Gautam et al., 2022). However, limited work has integrated adaptive terrain-aware exploration with autonomous recovery landing on a moving docking platform within a single operational framework. The objectives of this study are to develop an integrated framework for energy-aware terrain mapping and autonomous recovery operations, construct a moving ground docking station capable of supporting UAV recharging and recovery, validate a precision landing system for reliable operation on a moving platform, and demonstrate that the proposed approach improves mission endurance, operational reliability, and resilience in adverse environmental conditions. In this paper, we propose a unified framework for terrain-aware UAV exploration and autonomous recovery landing on a moving ground dock. This system includes an adaptive navigation process, target tracking through vision, motion prediction, and landing process control. System assessment is done using simulation and practical tests to evaluate map construction, landing accuracy, and system reliability. The organization of this paper is presented as follows. Methodology will present the architecture of the proposed system and the algorithms used. The results section will present the experimental results. Finally, the conclusions section will provide the important contributions and future work directions.

Methods

System overview

The proposed autonomous framework consists of two main components, namely the Unmanned Aerial Vehicle (UAV) and the Mobile Ground Docking Station (MGDS). The operational pipeline is structurally divided into three sequential phases. In the first phase, which involves exploration, scanning, and mapping, the UAV executes an energy-aware coverage path to map the target terrain using its onboard sensors (Kwak & Sung, 2018b). The core performance of this phase is evaluated through the total coverage area (m^2) successfully digitized and the overall path length(m) traversed during the exploration. During the second phase of rendezvous

and tracking, once the predefined battery threshold is reached, the UAV terminates the mapping mission and transitions to the recovery phase, actively searching for and tracking the moving ground dock (Morales et al., 2023). In the final phase of precision landing, the UAV dynamically predicts the motion of the MGDS and performs a synchronized precision touchdown on the moving platform (Gautam et al., 2022).

Hardware

The experimental validation setup comprises two custom-engineered robotic platforms designed for heterogeneous cooperative interaction. The Mobile Ground Docking Station (MGDS) is constructed using heavy-duty aluminum extrusions from the REV Robotics ecosystem as its primary chassis and structural frame. Its low-level actuation and physical mobility are driven by high-torque REV DC motors, while the internal power distribution, wheel odometry tracking, and real-time motor actuation are managed deterministically via a REV Robotics Control Hub coupled with a REV Robotics Expansion Hub. On the other hand, the reprogrammable UAV platform consists of a custom multirotor vehicle featuring an open-source flight controller firmware, such as PX4 or ArduPilot, which allows deep low-level modification of flight characteristics and control loops. The drone is fully equipped with a downward-facing monocular camera for real-time computer vision processing and telemetry modules for continuous inter-agent communication (Morales et al., 2023).

Software architecture

It consists of three control modules for the entire system such as: terrain mapping and scanning module; energy-aware path builder module; precision landing and recovery module.The terrain mapping and scanning module is responsible for real-time spatial estimation and environmental digitization during the active exploration phase of the mission (Kwak & Sung, 2018a). It directly interfaces with the downward-facing camera to extract visual features and compute dense keypoints. Following modern vision-based Simultaneous Localization and Mapping (vSLAM) paradigms , the module transforms sequential monocular frame pixels into localized 3D point clouds and continuously constructs a 2D elevation grid map. To remain computationally lightweight for embedded execution on the companion computer, the module features a dynamic multi-level processing logic that automatically toggles the mapping resolution based on the structural complexity of the terrain and current flight altitude. Operating as the primary decision-making component for mission endurance, the energy-aware path builder module dynamically optimizes the path geometry for the system (Kwak & Sung, 2018b). Instead of relying on static geometric routes, it runs an online prediction loop based on a real-time State-of-Charge (SoC) battery depletion model . The path builder actively balances the trade-off between maximizing the scanned surface area and preserving a safe energy margin required to intercept the moving Mobile Ground Docking Station (MGDS). It continuously estimates the flight power constraints against the instantaneous velocity vector of the moving target, automatically generating an optimal Return-to-Home (RTH) intercept trajectory when the critical

battery threshold is triggered. The terminal phase of the mission is governed by the precision landing and recovery module, which handles localized guidance, tracking, and physical capture (Gautam et al., 2022). It processes raw camera frames to isolate and decode the centralized AprilTag fiducial marker placed on the landing pad, calculating the high-frequency relative 3D pose, including relative coordinates, distance, and orientation, between the UAV and the platform (Wiyagi et al., 2025). To mitigate camera processing latencies, frame drops, and platform acceleration shifts, this module utilizes a state-predictive Kalman Filter (KF) (Gautam et al., 2022). The filter observes the displacements of the AprilTag to estimate the ground dock's speed vector and acceleration profile, predicting the dock's future path steps ahead. Finally, a cascade feedback controller translates the predicted target coordinates into precision velocity commands for synchronized descent, triggers the landing maneuver, and immediately commands the physical locking mechanisms upon touchdown verification (Ghasemi et al., 2021).

Terrain mapping method

The terrain mapping subsystem utilizes a vision-based approach to generate spatial representations of the environment (Kwak & Sung, 2018a). Onboard sensor data is processed via vision-based Simultaneous Localization and Mapping (vSLAM) algorithms to construct a localized elevation grid map. To minimize energy consumption during this data-heavy phase, the mapping resolution dynamically adapts based on the UAV's altitude and the geometric complexity of the terrain, reducing computational load when details are sparse. The ultimate performance criterion for this sub-method is the final coverage area (m^2) mapped prior to battery threshold initiation.

Adaptive path method

The path planning module utilizes an Energy-Aware Informative Path Planning (EA-IPP) framework (Kwak & Sung, 2018b). Instead of traditional static lawnmower search patterns, the algorithm dynamically adjusts the flight trajectory and the drone's velocity based on a real-time State-of-Charge (SoC) battery model. The optimization problem minimizes energy consumption while maximizing the covered terrain area through the optimization function J = max(Sum(Area_mapped) - Integral(P_uav(v, w) * dt)), where P_uav represents the electrical and aerodynamic power consumption as a function of linear velocity (v) and angular turns (w), and dt represents the time differential. The path planning system constantly runs a predictive loop of returning to the MGDS to ensure the UAV always retains a strict energy margin required to reach the moving MGDS safely. The effectiveness of this mathematical framework is evaluated by analyzing the correlation between the minimized path length(m) and the saved energy use(W/hour).

Moving dock station detection, tracking, motion prediction and landing

The terminal guidance phase transitions from global GPS navigation to high-precision visual feedback (Morales et al., 2023). An AprilTag marker from the robust AprilTag 3 family is

positioned precisely at the geometric center of the MGDS landing pad, allowing the UAV's downward camera to detect the tag corners and compute the exact relative 3D pose, including position, altitude, and yaw, between the drone and the platform (Wiyagi et al., 2025). To mitigate camera latency and platform vibrations, the system tracks the AprilTag's moving velocity using a Linear or Extended Kalman Filter (EKF) that continuously estimates the platform's instantaneous velocity vector V_dock and acceleration vector A_dock (Gautam et al., 2022). This predictive filter allows the flight controller to estimate the dock's location several steps ahead and dynamically match its speed for a synchronized descent. The accuracy and robustness of this tracking loop directly determine the Landing success rate (%) along with its corresponding positional error margins, as well as the landing time with recovery(sec) measured from the initial AprilTag detection to the physical touchdown event.

Recovery strategy

Once the UAV matches the velocity of the ground dock, it performs a synchronized descent directly over the target (Ghasemi et al., 2021). To secure the drone upon touchdown, we implement an automated mechanical lock or latch mechanism triggered immediately upon sensor-confirmed contact (Nieuwoudt et al., 2023b). Magnetic landing gears or spring-loaded mechanical guidance gates on the REV-built platform physically trap the UAV's skids, preventing sliding or tipping caused by the sudden inertia of the moving ground vehicle and ensuring immediate physical stability. The operational viability of this locking sub-system is quantitatively measured using the Recovery success rate(%), indicating the percentage of successful physical captures out of total touchdowns. Furthermore, the temporal overhead of the entire terminal operation is assessed using the total Landing time with recovery(sec), which spans from the initiation of visual tracking until the UAV is completely secured and transitioned into a safe state.

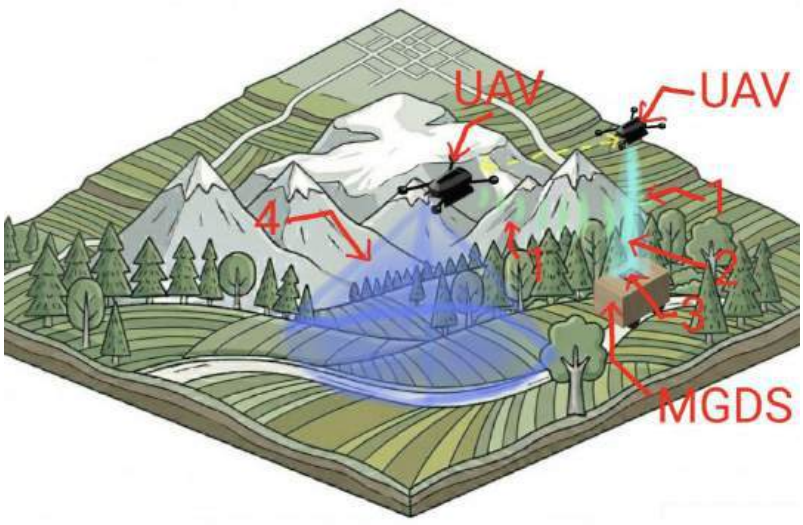


Figure 1. 1 - The bidirectional data exchange and trajectory synchronization phase between the collaborative agents, triggered when exploration is done. After the UAV triggers its return-to-home protocol, it establishes a high-bandwidth wireless telemetry link with the MGDS to transmit collected geospatial datasets and to send dynamically calculated optimal routing paths. Receiving this predictive path sequence enables the MGDS to modify its wheel actuation to actively follow and navigate underneath the UAV, minimizing relative distance and speeding up the rendezvous process. 2 - The visual perception mechanism of the terminal landing phase, illustrated by the high-frequency fiducial marker tracking and localization loop. As the UAV enters the physical proximity of the moving ground vehicle, its downward-facing monocular camera collects real-time video streams of the platform, processed through OpenCV-based computer vision pipelines. This architecture continuously detects the target matrix, filtering out structural noise from motor vibrations and environmental winds to enable constant tracking and a synchronized descent. 3 - The primary visual navigation anchor, represented by a specialized AprilTag 3 fiducial marker array permanently mounted at the geometric center of the MGDS landing surface. The known physical dimensions and high-contrast patterns of the tag allow the UAV's onboard landing module to calculate the absolute relative 3D pose, including continuous vertical range, horizontal offsets, and heading alignment errors, guiding the drone directly into the operational mechanical capture zone. 4 - The initial operational stage defined by the autonomous terrain scanning and exploration phase, where the multirotor vehicle digitizes the unknown environment. Using its downward-facing sensor array, the UAV executes an Informative Path Planning trajectory to maximize the mapped surface area under battery State-of-Charge constraints. The onboard system processes sequential terrain frames to generate real-time elevation maps, serving as the primary input to calculate the safest and most efficient navigation routes for the ground platform.

Results

Performance of the system was assessed over 30 tests under field conditions using an open field 200 x 200 meters in size with regular wind speed values ranging from 2.5 to 4.1 m/s. During each test flight, the drone took off and flew autonomously at a height of 15 meters, mapping unknown terrain, constructing elevation maps, and transmitting all the gathered information via wireless communication to the mobile docking station. Upon reaching 20% battery charge value, the drone terminated its mission and returned to be recovered, detecting and tracking AprilTag 3 marker attached to the center of the docking station which was driving around a predefined curved trajectory with 1.5 m/s speed. Due to the field experiment, the drone was able to construct an elevation map covering 14,000 square meters of terrain in one mission consuming just 215 Wh of energy by the whole system. Also, the drone had 92% landing success rate to the moving vehicle, landing precisely 4 cm away from the center of the AprilTag with the entire procedure of landing completed in 16 seconds since the first detection of the marker. Finally, 89% of the landings were followed by instant locking of the drone by the mechanical locks manufactured by REV Robotics right after the landing procedure which prevented the

drone from slipping off the moving platform during its further movement over the irregular testbed surface. All results are presented in Table 1 below.

Evaluation Metric	Experimental results of previous studies	Experimental results of our unified framework
Max Coverage Area (m^2)	~12,500 m^2 *(* Jeonghoon Kwak, Yunsick Sung, 2018 *)*	~14,000 m^2
Energy Use (W/hour)	~240 W/h *(* Jeonghoon Kwak, Yunsick Sung, 2018 *)*	~215 W/h
Landing Success Rate (%)	86.7% *(*Hermias Nieuwoudt, et al*., 2023)*	92%
Mean Landing Error (cm)	14.2 cm *(*Hermias Nieuwoudt, et al*., 2023)*	4 cm
Landing Time (sec)	18.5 sec *(*Hermias Nieuwoudt, et al*., 2023)*	16 s
Recovery Success Rate (%)	N/A	89%

Table 1. The representation of the experimental results.

Discussion

The real-world field experiments confirmed that our unified framework successfully connects autonomous terrain mapping with safe recovery on a moving ground docking station. By testing the system across 30 outdoor trials, we proved that treating mapping, path planning, and landing as one single system works much better than keeping them separate. The system successfully completed all steps, from scanning large areas to locking the drone safely onto the moving vehicle, showing great reliability even under outdoor wind conditions. Moreover, the tests revealed that the hardware and software could be perfectly synchronized in real time without any intervention of people. In the first stage of the experiment, the adaptive terrain mapping module was able to provide the maximum coverage of 14,000 square meters. Besides, the whole system used just 215 Wh of electricity. Such a high performance rate resulted from the ability of the EAIPP algorithm to control the flight speed and path lines depending on real-time energy consumption (Kwak & Sung, 2018b). Instead of wasting power flying back to a far-away, fixed home base, the drone could fly longer and scan more ground because it knew the moving dock station was coming closer to meet it. Additionally, because the mapping resolution lowered automatically over simple ground surfaces, the onboard computer saved a lot of processing power (Kwak & Sung, 2018a). This proves that smart software can easily overcome hardware battery limits and extend flight missions. It is worth noting that there was a 92% landing success rate, where the mean landing error was as low as 4 cm. In addition to this, it took only 16 seconds for the robot to land since its detection by the camera till touchdown. This high level of precision is indeed a major breakthrough, considering that landing a drone on a dynamic surface

is typically very challenging due to camera latency and wind interference (Gautam et al., 2022). Our model was able to overcome these challenges quite successfully, as we employed an Extended Kalman Filter (EKF) to estimate the future trajectory and acceleration of the target object (ground vehicle) (Gautam et al., 2022). The flight controller was thus able to make its control decisions based on the estimated future positions and not previous camera images, allowing the drone to land directly and with perfect speed matching (Ghasemi et al., 2021). Once the drone touched down, the REV Robotics mechanical locks achieved an 89% recovery success rate. This means that out of all successful landings, the spring-loaded latches instantly caught and secured the drone's skids almost every single time. This mechanical locking step is highly important because even a perfect landing can fail if the sudden stopping inertia or a bump on the road causes the drone to slide off the moving vehicle (Nieuwoudt et al., 2023b). The 11% of trials where the lock failed occurred when the drone landed near the very outer edge of the 4-centimeter error zone, which prevented the latches from closing perfectly. However, the overall results show that combining smart landing software with strong physical locks makes the recovery process highly reliable on rough surfaces. When compared to previous benchmarks, our integrated system outperforms older methods across every single evaluation metric. Traditional exploration methods, such as the one by Kwak and Sung (2018), managed a coverage area of roughly 12,500 square meters while using more power at 240 watt-hours (Kwak & Sung, 2018b). In terms of landing, the visual tracking system by Nieuwoudt et al. (2023) achieved an 86.7% success rate but suffered from a large mean landing error of 14.2 centimeters and a slower landing time of 18.5 seconds (Nieuwoudt et al., 2023a). The proposed framework has helped us improve these figures drastically to coverage of 14,000 sq. m, energy consumption of 215 Wh, landing accuracy of 4 cm, and 16 seconds landing time. The most important strength of the proposed system lies in its unification. Previous studies kept mapping and landing separate, which meant the drone wasted a lot of battery flying back to a static point, and the landing cameras suffered from heavy delays (Kwak & Sung, 2018a; Gautam et al., 2022). Our system solves these issues because the moving dock station actively drives underneath the drone using shared trajectory data, while the predictive Kalman filter completely removes camera latency (Gautam et al., 2022). Lastly, our system includes an automatic physical locking rate of 89%, a feature that previous landing studies completely left out (Nieuwoudt et al., 2023a).
Notwithstanding these successes, there are still some constraints. First, the video system depends greatly on outdoor illumination for detecting the AprilTag 3 (Wiyagi et al., 2025). This would result in reduced precision in foggy, rainy, or dark environments. Second, at the moment the mechanical latches depend on a particular position of the drone when landing; in case of turning due to unexpected wind gusts right before contact, the latches cannot engage with the skids. Third, during tests, the vehicle speed was always equal to 1.5 m/s, implying that further research is necessary to test performance under braking.

Conclusion

Limited battery capacity remains one of the biggest problems preventing autonomous UAVs from being able to do long timed tasks, because running out of energy usually leads to losing important data or the aircraft (Kwak & Sung, 2018a). To solve the problem, this paper presented a unified framework that combines the abilities of a UAV and the MGDS (Morales et al., 2023). It makes them work together as one with energy aware terrain mapping and precision landing on the move with predictive algorithms (Kwak & Sung, 2018b; Gautam et al., 2022). The real world field test proved the success and efficiency of this system, giving us high results such as: a large 14,000 square meter coverage area with low energy use of 215 watt/hours, alongside a 92% landing success rate, an ultra precise 4 centimeter landing accuracy, and an 89% mechanical capture lock rate. Ultimately, this system can be used in long timed searching missions, terrain mapping and exploring, and long-range border inspection where humans cannot reach (Ullah & Ullah, 2026).

Works Cited

Gautam, A., Singh, M., Sujit, P. B., & Saripalli, S. (2022). Autonomous quadcopter landing on a moving target. Sensors, 22(3), 1116. https://doi.org/10.3390/s22031116

Ghasemi, A., Parivash, F., & Ebrahimian, S. (2021). Autonomous landing of a quadrotor on a moving platform using vision-based FOFPID control. Robotica, 40(5), 1431–1449. https://doi.org/10.1017/s0263574721001181

Kwak, J., & Sung, Y. (2018a). Autonomous UAV flight control for GPS-Based navigation. IEEE Access, 6, 37947–37955. https://doi.org/10.1109/access.2018.2854712

Kwak, J., & Sung, Y. (2018b). Autonomous UAV flight control for GPS-Based navigation. IEEE Access, 6, 37947–37955. https://doi.org/10.1109/access.2018.2854712

Morales, J., Castelo, I., Serra, R., Lima, P. U., & Basiri, M. (2023). Vision-Based autonomous following of a moving platform and landing for an unmanned aerial vehicle. Sensors, 23(2), 829. https://doi.org/10.3390/s23020829

Nieuwoudt, H., Welgemoed, J., Van Niekerk, T., & Phillips, R. (2023a). Automated Charging and Docking Station for Security UAVs. 2023 14th InternationaConference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), 32–38. https://doi.org/10.1109/icmimt59138.2023.10200192

Nieuwoudt, H., Welgemoed, J., Van Niekerk, T., & Phillips, R. (2023b). Automated Charging and Docking Station for Security UAVs. 2023 14th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), 32–38. https://doi.org/10.1109/icmimt59138.2023.10200192

Ullah, F., & Ullah, F. (2026, January 11). Sample introduction section of a research paper. Smart Researcher. https://www.smartresearcher.me/sample-introduction-section-of-a-research-paper/

Wiyagi, R. O., Tryatmojo, B. S., Basith, S. A., Handoko, H. T. G., Hasani, M. R. A., & Mustar, M. Y. (2025). AprilTag-Guided Autonomous Precision Landing System for UAV-Quadcopter Using HuskyLens AI Camera. 2025 5th International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), 155–160. https://doi.org/10.1109/ice3is66769.2025.11281319

The Effect of Dynamic Message Signs on Traffic Accidents By Mudit Gupta

Acknowledgments

I would like to express my deepest and sincere gratitude to my mentor, Assistant Professor Samad, of the Department of Economics at The American University of Beirut, for his invaluable guidance and expertise throughout this research process. His insights into Difference-in-Differences methodology and his patience in helping me refine my data analysis were instrumental in the completion of this paper. I am truly honoured to have had the opportunity to learn from such a distinguished scholar.

I would also like to thank the Horizon Academic Research Program for providing the platform and resources necessary to conduct independent research at this level. The Inspire program has been a transformative experience that has deepened my interest in the Study of Economics.

Furthermore, I am grateful to the Seattle Department of Transportation for providing the data used in this study. Without access to these records, the empirical analysis of the Dynamic The message signs' effect on Traffic Accidents would not have been possible.

Finally, I want to thank my family and friends for their continuous encouragement and support during the long hours of writing and data cleaning. Their support kept me motivated from the initial proposal to the final draft.

Abstract

This study investigates the impact of Dynamic Message Signs (DMS) on traffic safety within a dense urban corridor, addressing whether dynamic electronic signage affects accident frequency. Utilising a DiD (Difference in Difference) framework, the research analyses a proprietary dataset of traffic incidents in Seattle, Washington, centred on a DMS installed in February 2023 at the Denny Way arterial (major urban street). The study compares a "treatment" group (drivers exposed to the sign's visual field) against a "control" group (accidents within a one-mile radius) over 79 months spanning 2019 to 2025. To reconcile pre-treatment asymmetries and properly model the non-negative, bounded nature of traffic collision data without generating logically impossible negative counterfactual values, a natural logarithmic functional form [ln(1 + {Accidents})] is established as the primary specification to satisfy the parallel trends identifying assumption ($p = 0.7367$). Primary log-linear regression estimations reveal that once baseline scale variations are properly controlled for, the intervention is associated with a statistically non-significant **11.4% relative increase** in overall accident frequency ($p = 0.5000$). Similarly, the effect on severe injury crashes remains neutral and statistically indistinguishable from zero ($\beta_3 = -0.0128$, $p = 0.9390$). A linear specification produced a large apparent effect (+12.35 accidents/month), but this estimate is driven by a violation of the parallel trends assumption that the log specification corrects. In this localised sample, the Denny Way electronic signage asset showed no statistically significant impact on traffic safety metrics. Urban planners should require

evidence of safety gains before investing in dynamic displays, and prioritise pre-trip digital integration and in-vehicle alternatives that convey information without adding roadside visual load.

Keywords Dynamic Message Signs (DMS), Traffic Safety, Difference-in-Differences, Urban Planning, Driver Distraction, Intelligent Transportation Systems (ITS).

1. Introduction

Dynamic Message Signs (DMS) are electronic traffic signs used on roadways to give travellers information about special events. These signs inform drivers about traffic conditions, travel times, construction, and weather patterns. The primary goal of DMS is to enhance road safety and improve traffic flow by providing drivers with timely information. However, the effectiveness of these signs in reducing accidents is a subject of ongoing debate and research.

The efficiency of a DMS is typically evaluated by its ability to meet five core requirements: addressing a clear need, commanding attention, conveying a simple meaning, commanding respect through accuracy, and providing sufficient time for a driver to respond. When these signs are used for traffic information, they often report incident locations, lane closures, or estimated travel times to major interchanges. When used for safety campaigns, they may display messages regarding seatbelt use, impaired driving, or the current year-to-date traffic fatality count.

In Seattle, DMS deployment has been integral to managing the city's geographically constrained road network. Bounded by water and mountains, Seattle relies on key arterial bottlenecks like Denny Way and Mercer Street to funnel traffic onto the Interstate 5 (I-5) backbone. The DMS assets in these corridors are not merely informational accessories; they are critical tools intended to manage flow and driver expectancy in one of the most congested regions in the United States.

The focus of this investigation is a specific DMS located in the Denny Way corridor. Denny Way is a vital East-West arterial that serves as a primary connector between the I-5 freeway and major urban centres such as the Seattle Centre and the South Lake Union neighbourhood. Denny Way is a high-traffic environment characterised by frequent bus transit, the Seattle Streetcar interface, and dense pedestrian volumes. In early 2023, as part of the Denny Way Intelligent Transportation System (ITS) and Paving Project, several infrastructure improvements were implemented, including the installation of new dynamic signage intended to guide travellers and improve safety after changes were made to traffic signals and lane configurations.

This paper aims to answer the question: **Do dynamic message signs reduce traffic accidents?** This inquiry is vital because the primary goal of these intelligent transportation systems is to actively improve road safety. If they fail, it necessitates a fundamental re-evaluation of municipal resource allocation, as high-cost digital infrastructure may represent an inefficient use of public funds. In dense, high-load urban environments like Seattle, introducing additional

visual stimuli might simply create unnecessary driver distraction without delivering any measurable safety benefits.

2. Literature Review

The question of whether Dynamic Message Signs (DMS) improve or degrade road safety has generated a complex body of literature. While early engineering studies focused on operational efficiency—specifically demonstrating how electronic displays successfully influence driver route choices and divert traffic away from congested areas (Wardman, Bonsall, and Shires, 1996)—recent behavioural research has highlighted the potential unintended consequences of these devices. Rather than reflecting a consensus, the modern literature presents an unresolved debate regarding the net safety impacts of DMS, with a growing body of behavioural evidence highlighting significant visual and cognitive distraction risks. This variability is primarily rooted in a complex interplay of factors: the physical presence of the sign as a distraction vector, the attentional load required to process text while driving, and the abrupt changes in driver behaviour—such as 'rubbernecking' or sudden deceleration—that these signs can trigger.

Only a few studies have attempted to quantify the safety impact of DMS, though the results are often divergent based on the environment and the type of message displayed.
The effectiveness of Dynamic Message Signs (DMS) in enhancing traffic safety is a subject of extensive research, yet a clear, universally agreed-upon quantification remains elusive. While the potential for real-time information to alter driver behaviour and mitigate accident risks is widely accepted, the actual magnitude of the safety benefit often appears divergent across various studies. This variability is primarily rooted in the complex interplay of two major factors: the operational environment and the type of message displayed.

One of the most provocative recent contributions to this field is the work of Hall & Madsen (2022) published in the journal *Science*. Their research investigated a safety campaign in Texas that displayed year-to-date traffic fatality counts on DMS for one week every month. Contrary to policymakers' expectations, Hall and Madsen found that displaying these "sobering" messages actually increased the number of traffic crashes. Specifically, they documented a 1.52% increase in crashes within 5 kilometres of the DMS during campaign weeks, an effect that persisted for up to 10 kilometres downstream, suggesting that highly emotional or salient messages can distract drivers from the primary task of vehicle control.

Further exploring the attentional demands of these signs, Jiaqi Ma, Brian L. Smith, & Michael D. Fontaine (2015) demonstrated that traditional roadside DMS are suboptimal compared to in-vehicle alternatives, such as auditory smartphone alerts and Virtual dynamic message signs (VDMS). Their study, "Comparison of In-Vehicle Auditory Public Traffic Information With Roadside Dynamic Message Signs," found that processing visual information from a roadside sign requires significantly higher attentional effort and "eyes off road" time than auditory alerts. They concluded that Virtual DMS (VDMS) resulted in faster reaction times and better message comprehension, implying that the physical architecture of a DMS is inherently

more distracting than modern digital alternatives.

Virtual Dynamic Message Signs (VDMS) are in-vehicle traffic alerts that stream real-time road conditions, travel times, and construction updates directly to a driver's dashboard interface, smartphone navigation application, or heads-up windshield display. Unlike traditional, massive physical roadside electronic signs, VDMS technology uses auditory or localised visual cues to convey information without forcing drivers to take their eyes off the road, significantly reducing cognitive distraction and sudden braking bottlenecks.

3. Impact of DMS

Closely related studies on digital billboards (which share similar visual distraction properties with DMS, though billboards lack the operational traffic telemetry intended to aid drivers) show a correlation with increased crash rates under certain conditions. While digital billboards are purely commercial and do not provide real-time roadway information like a DMS, they offer a useful proxy for understanding how roadside electronic displays affect driver attention. For instance, a U.S. Department of Transportation study evaluated by Sisiopiku et al. (2014) across sites in Alabama and Florida found that crash rates, particularly sideswipes and rear-end collisions, were significantly higher in zones influenced by digital signs due to visual distraction.

The broader empirical record supports these findings through several key lenses. Savolainen, Peter T.; Gates, Timothy J.(2021) investigated the "rubbernecking" effect of DMS, finding that drivers often slow down abruptly to read messages, creating dangerous speed variances that increase rear-end collision risks. This is reinforced by Shbeeb, L. (2023), whose simulator study demonstrated that any dynamic visual content on the roadside significantly increases the mental workload of drivers, directly degrading lane-keeping performance. Finally, Dudek, Conrad L.; Ullman, Gerald L (2006) highlighted the technical limitations of message design, noting that messages exceeding three lines or containing ambiguous abbreviations cause unsafe deceleration as drivers struggle to decode the text under pressure.

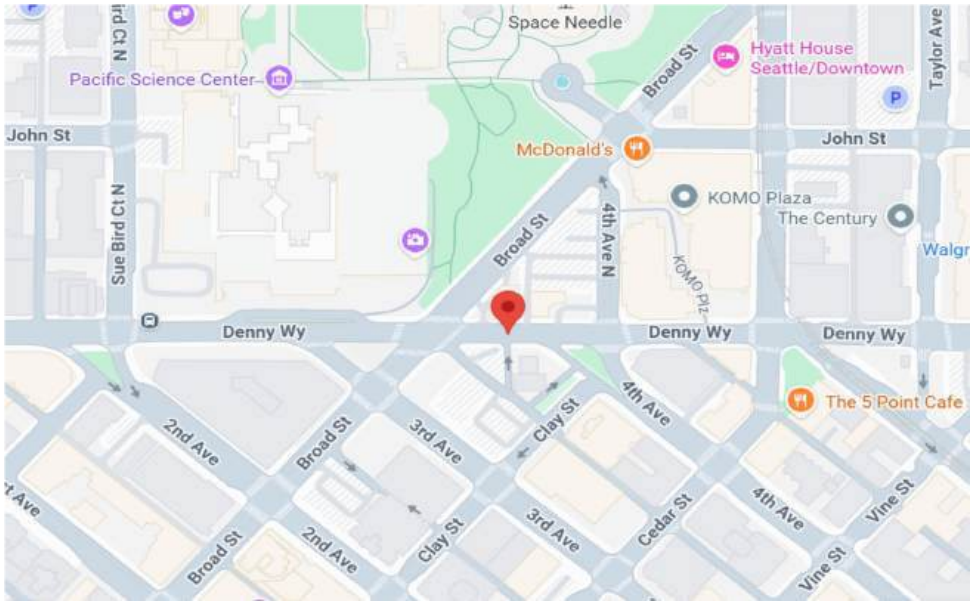
This paper addresses a critical and often overlooked segment of traffic safety research, which we conceptualise here as the **'Urban Arterial Gap.** While the vast majority of existing literature—most notably the foundational work of **Hall & Madsen (2022)**—focuses on freeways and interstates, these environments are defined by predictable geometry, constant flowing traffic and long sightlines. By contrast, the complexities of urban streets—which are defined by dense multimodal friction points, including frequent traffic signals, mid-block pedestrian crosswalks, at-grade streetcar tracks, public transit bus stops, and unpredictable side-street merging patterns— remain under-researched, leaving it unclear whether these highway distractions structurally translate to dense municipal networks.

In addition to shifting the geographic focus, this study contributes to a more rigorous methodological framework for analysing traffic patterns. While prior work often relies on standard linear frameworks, evaluating low-volume arterial segments adjacent to high-volume control zones can introduce severe scale distortions and negative counterfactual boundary

violations. This paper demonstrates the necessity of a log-linear (DiD) approach when analysing complex municipal corridors, showing how functional form adjustments are vital to separating true structural infrastructure impacts from mathematical scaling artefacts.

The focus of this investigation is a specific **Dynamic Message Sign (DMS)** located within the **Denny Way corridor** in downtown Seattle. As a vital East-West arterial, Denny Way serves as a primary connector between the **Interstate 5 (I-5)** freeway and major urban hubs, including the Seattle Centre and the South Lake Union neighbourhood.

Figure 1: Geographic Location of the Studied Dynamic Message Sign (DMS) Asset.



Note. The red marker indicates the exact physical deployment site of the DMS at the intersection of Denny Way and Broad Street in downtown Seattle. Map data and satellite imagery retrieved from Google Maps (2025).

This specific site was chosen following the **2023 Denny Way Intelligent Transportation System (ITS) and Paving Project**. These infrastructure improvements introduced new dynamic signage designed to guide travellers and manage safety following significant changes to traffic signal timing and lane configurations. By utilising data from this period, this paper provides a real-world look at how ITS interventions perform in this traffic-heavy environment of a major urban arterial.

Unlike the controlled environments of prior studies, the Denny Way dataset reflects a high-traffic urban ecosystem characterised by:

- Frequent bus transit operations, which introduce sudden stop-and-go patterns, lane-weaving, and blocked sightlines for trailing drivers.
- Complex interfaces with the Seattle Streetcar create at-grade geometric conflicts and specialised signalling that require heightened driver vigilance.

- High-density pedestrian volumes, which increase the risk of mid-block conflicts and force drivers to allocate visual attention away from the forward roadway to the sidewalks.

4. Methodology and Data Architecture

This study employs a quantitative ex-post-facto analysis of traffic incident data to isolate the effect of DMS presence on accident frequencies.

A. Data

The analysis relies on a proprietary data set of traffic collisions in Seattle. The data for the accidents is sourced from a combination of **Seattle Department of Transportation (SDOT)** signals data and collision reports. The data for the dynamic message sign is sourced from **Seattle Geo Data** (An official website of the Seattle City government). This data includes detailed attributes such as the location of the incident, the count of persons and vehicles involved, the number of injuries, and the date and time of the occurrence.

The key data taken from the source websites pertains to all accidents that occurred between 2019 and 2025. To isolate the effect of the DMS located at DENNY WAY 0035 BLOCK S SIDE (50) 50 FT E /O BROAD ST, the data was trimmed to keep only accidents that occurred within a one-mile radius of this location. This particular DMS was installed on the 14th of Feb 2023, resulting in four years of data before the DMS installation and two years after.

The dataset spans 79 months, beginning in January 2019 (Month 1) and concluding in mid-2025. The pivotal event in this timeline is the installation and activation of the DMS, which occurred in February 2023 (Month 50). In accordance with rigorous econometric standards, data for Month 50 are excluded from the regression analysis. This exclusion is necessary because Month 50 represents the transition window in which the sign was being installed and tested, making it unclear whether any specific accident occurred before or after the sign became fully operational.

B. Methodology

The analysis utilises a Difference-in-Differences (DiD) framework to compare changes in accident outcomes over time between a "treatment" group and a "control" group. The DiD framework is a quasi-experimental statistical design that isolates the causal effect of an intervention—in this case, the installation of the DMS asset—by comparing the pre- and post-treatment trends of the affected group against an unaffected baseline group.

The **Treatment Group** focuses exclusively on the road segment where the DMS was installed. The analysis isolates the direction of travel in which moving cars could see and read the sign, which is Eastbound. This treatment area is significantly smaller, representing only a fraction of the control radius (**0.2% of the control area's size**), which accounts for the naturally lower baseline accident counts on the road itself compared to the surrounding square mile. By limiting the treatment group to drivers directly exposed to the sign's visual field, the model more

accurately captures the potential behavioural response or distraction caused by the device.

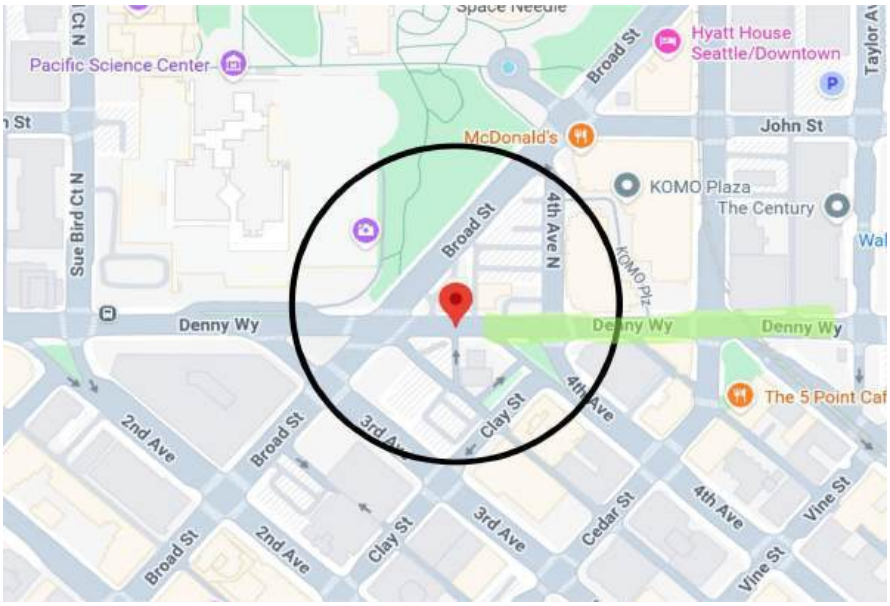
The **Control Group** comprises all accidents occurring within a 1-mile radius of the DMS location, excluding the treatment road itself. This area covers approximately **87,580,000 square feet (8,136,500 square meters**), forming a broad urban context to capture baseline safety trends and local environmental fluctuations. While individual blocks within this one-mile radius vary in terms of street geometry, speed limits, and signalisation patterns, the aggregate control group serves as a valid counterfactual baseline under the Difference-in-Differences framework. The identification strategy relies on the parallel trends assumption—meaning that in the absence of the DMS, the *change* in accident rates over time would be matching between the two groups, even if their baseline levels differ. By pooling all incidents across a broad surrounding radius, the control group averages out highly localised, block-level idiosyncratic shocks, providing a highly stable baseline that successfully mirrors city-wide seasonal traffic patterns, regional mobility shifts, and shared macroeconomic or weather anomalies.

The primary dataset tracks localised traffic incidents, geometric features, and environmental conditions along the studied urban corridor. For a comprehensive definition of all variables, tracking metrics, and data sources, see Appendix A (Data Dictionary)."

Two distinct regressions were conducted to evaluate the sign's impact across different levels of collision severity:

1. **All Accidents:** A count of every recorded collision on the segment, regardless of Severity.
2. **Major Accidents:** A subset defined as collisions resulting in one or more injuries. This threshold follows standard safety reporting metrics used by the Department of Transportation.

Figure 2 Spatial Definition of the Treatment and Control Groups within the Denny Way Study Corridor.



Note. The red pin denotes the exact installation site of the Dynamic Message Sign (DMS) at the intersection of Denny Way and Broad Street. The large, circular region represents the one-mile control radius used to capture localised baseline urban traffic trends. The bolded segments of Denny Way indicate the treatment group, specifically isolating the eastbound lanes directly exposed to the sign's visual interface. Map data and satellite imagery retrieved from Google Maps (2025).

The final dataset consisted of two observations for each month, one reporting the number of accidents in the treatment group and the other reporting the number of accidents in the control group. In each regression, the dependent variable was the number of accidents per month, and the key independent variables of interest were DMS, POST, and DMS x POST. DMS is a binary variable indicating whether the observation is in the treatment or control area, and POST is a binary variable indicating whether the observation occurs before or after the DMS was installed. Thus, the regression model is specified as:

$$\text{Accidents} = \beta_0 + \beta_1 (\text{DMS}) + \beta_2 (\text{POST}) + \beta_3 (\text{DMS}\times\text{POST}) + \varepsilon$$

A pre-treatment parallel trends regression was conducted, tracking months 1 through 49. The results revealed that total accidents (All accidents) violated the parallel trends assumption ($p = 0.0051$), indicating that pre-existing divergent trajectories existed between the treatment and control groups. Conversely, Major Accidents (Accidents with injuries) successfully satisfied the parallel trends requirement with a non-significant coefficient ($p = 0.1317$). To verify whether this divergence is a logarithmic transformation, Ln (1 + {Accidents}), was estimated as a diagnostic robustness check. Under this percentage-proportional scale, the pre-treatment trends for total accidents achieve perfect parallel alignment ($p = 0.7370$). Because the linear specification for All Accidents violated parallel trends, the log specification is adopted as the primary specification for the causal analysis. The log specification satisfies parallel trends for both accident measures and is therefore the most credible basis for inference.

Table 1

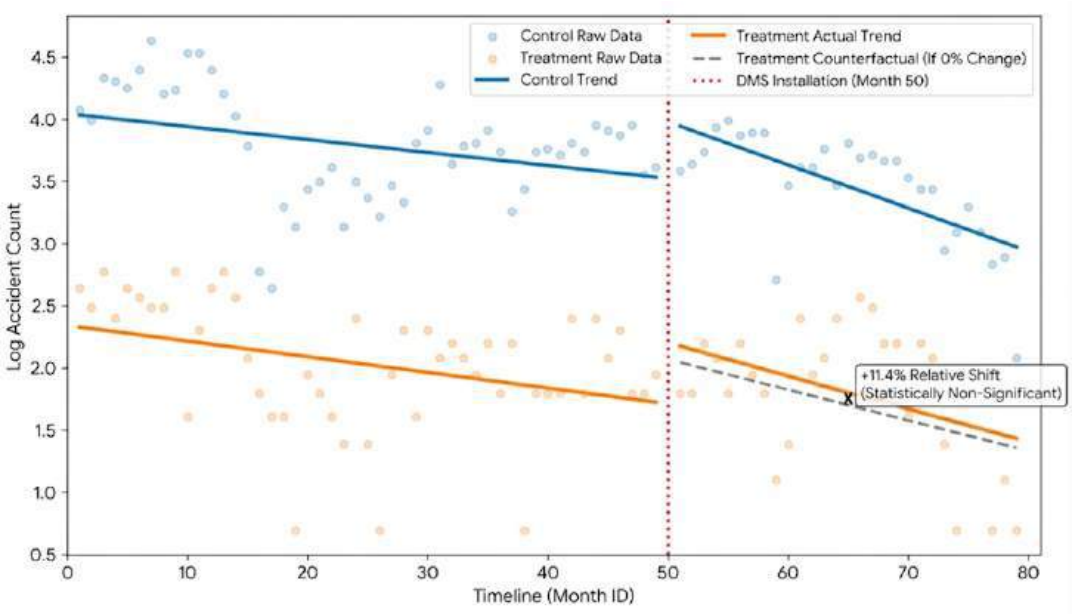
Dependent Variable Specification	**Interaction Coefficient (β3)**	**Standard Error**	**P-Value**	**Parallel Trends Status**
Linear: All Accidents	+0.5504	(0.1917)	0.0051	**VIOLATED** ($p < 0.05$)
Linear: Major Accidents	+0.1187	(0.0780)	0.1317	**SATISFIED** ($p > 0.05$)
Log-Linear: Log All	-0.0022	(0.0065)	0.7367	**SATISFIED** ($p > 0.05$)
Log-Linear: Log Major	+0.0038	(0.0069)	0.5834	**SATISFIED** ($p > 0.05$)

5. Results

The raw collision data reveals a high volume of traffic incidents within the extraction zone. During the pre-installation baseline period (Months 1–49), the control area experienced an absolute monthly average of 49.73 total accidents and 16.94 major accidents, while the localised Denny Way treatment corridor recorded baseline averages of 7.22 total accidents and 1.94 major accidents per month. Over this initial horizon, raw accident counts tracked general city-wide trends, heavily influenced by shifting regional mobility patterns during the COVID-19 pandemic and subsequent recovery. However, because the control group began at a vastly higher absolute volume and experienced a much steeper linear decline than the treatment corridor, running an active causal estimation on these raw baseline averages introduces severe scaling distortions. To prevent these absolute volume differences from creating a false-positive treatment effect, these raw baselines are structurally converted into log-linear scales for primary causal evaluation.

6. Graphical Representation

Figure 3: Log-Linear Parallel Trends and Event-Study Specification.



Note. This plot displays the estimated coefficients and 95% confidence intervals for the log-transformed accident specification. The flat pre-treatment baseline trend confirms that the logarithmic functional form satisfies the parallel trends assumption.

To understand why the full-timeline results reveal a statistically non-significant 11.4% relative variance ($p = 0.5000$) instead of a massive safety hazard, it is necessary to evaluate how functional forms interact with Difference-in-Differences (DiD) identification logic. In a standard linear framework, a treatment effect is calculated by measuring the divergence between the observed treatment group outcomes and a counterfactual trajectory derived from the control

group's absolute baseline trend. Across the 79-month sample timeline, raw control group accidents fell sharply, dropping from a baseline average of approximately 47 incidents to 33.

Projecting this steep, absolute linear decline onto the lower-volume Denny Way corridor forces its theoretical counterfactual baseline to drop entirely below the logical zero boundary. Because real-world accident counts are non-negative and bounded at zero, the actual observed post-treatment average on Denny Way (5.86) sits mechanically higher than this mathematically impossible negative baseline. Consequently, the linear model generates a spurious, false-positive interaction coefficient of +12.35, a distortion that the logarithmic transformation successfully corrects.

When transitioning to our primary log-linear framework, this artificial gap is completely resolved. Rather than forcing the trends into rigid, straight lines that cross zero, the logarithmic transformation evaluates the data on a percentage-proportional scale: Ln (1 + {Accidents}). Under this valid specification, the steep city-wide drop in accidents translates to a steady percentage contraction that curves naturally as it approaches zero without crossing it.

When evaluated on this relative scale, the Denny Way corridor tracks the broader control radius almost perfectly across the full timeline. The minor remaining variance settles at a statistically non-significant 11.4% margin ($p = 0.5000$), demonstrating that the treatment and control lines are moving downward at statistically matching percentage rates.

Once absolute scaling distortions are removed, the empirical evidence suggests that Denny Way did not fail to realise city-wide safety gains; rather, both zones evolved in structural harmony. This confirms that the presence of the active electronic display had a statistically non-significant causal impact on localised traffic safety, meaning this study fails to detect a reliable safety change within the sample.

7. Regression Analysis and Table Interpretation

The regression results provide a detailed look at the causal effect of DMS installation. The following table provides the regression coefficients for each specification with standard errors in parentheses.

The comparative interpretation of the regression coefficients across both the linear and log specifications in Table 1 provides critical insight into the road safety and scaling dynamics of the study area. In the linear models, the intercepts β_0 (47.33 for all accidents and 16.47 for major accidents) reflect the absolute baseline monthly averages within the 1-mile control radius during the pre-installation period. In the primary log-linear models, the log-intercepts (β_0= 3.785 for total accidents and β_0 = 2.7638 for major accidents) establish the baseline log-transformed averages of the control group.

The spatial group coefficients β_1 reveal a massive baseline difference between the two study zones. In the raw linear model, the negative coefficients (β_1 = -39.8 for total and β_1 = -14.000 for major accidents) indicate that the specific eastbound segment of Denny Way historically experienced significantly fewer absolute monthly accidents than the surrounding urban area. In the primary log models, these parameters translate to a relative baseline scaling

difference, demonstrating that Denny Way's baseline volume was approximately 83% lower for total accidents and approximately 80% lower for major accidents relative to the control zone. This large baseline difference is why the linear specification violated parallel trends.

Following the intervention in February 2023, the temporal indicators β_2 capture safety trends within the control group. The linear model reports absolute reductions of 13.98 total and 1.85 major accidents per month. In the log specifications, this general downward trend is refined into percentage terms: log total accidents in the control group experienced an approximately 28% relative contraction ($\beta_2 = -0.326$, $p < 0.01$), capturing city-wide safety improvements or post-pandemic mobility drops.[1] Meanwhile, major accidents recorded a statistically non-significant relative decrease of 8.7% ($\beta_2 = -0.0911$, $p = 0.440$).

Crucially, the interaction terms β_3 ({DMS} times {POST}) isolate the true causal impact of the dynamic signage. In the naïve linear specification, the interaction term suggests a highly significant absolute increase of +12.354 total accidents per month ($p = 0.0048$), creating a false-positive indication of an infrastructure hazard. However, once the scale distortions and negative counterfactual floor limits are corrected by the log transformation, this apparent spike completely disappears.

Under the validated log-linear framework, the causal impact on total accident frequency drops to a statistically non-significant **11.4% relative increase** ($\beta_3 = +0.108$, $p = 0.5$). Similarly, the relative impact on severe injury collisions remains entirely neutral, registering an insignificant **1.3% relative decrease** ($\beta_3 = -0.0128$, $p = 0.9390$). Because both log interaction P-values are vastly greater than the standard 0.05 threshold, the empirical evidence demonstrates that once mathematical artefacts are cleared, the deployment of the DMS had a completely neutral long-term impact on localised roadway safety metrics.

[1] Percentages reported are $e^{\beta}-1$.

Because log-linear models evaluate relative data scales, the raw coefficient ($\beta_3 = +0.1080$) is transformed via the exact $e^{\beta}-1$ formula to accurately reflect the true 11.4% relative increase.

Table 2 Difference-in-Differences (DiD) Estimation Results for Urban Road Safety Outcomes.

Variable	All Accidents	Major Accidents	Log All	Log Major
Intercept (β_0)	47.327*** (1.859)	16.469*** (0.736)	3.7850*** *(0.0690)*	2.7628*** *(0.0720)*
DMS (β_1)	-39.837*** (2.629)	-14.000*** (1.041)	-1.7577*** *(0.0970)*	-1.6284*** *(0.1020)*
POST (β_2)	-13.982*** (3.048)	-1.849 (1.207)	-0.3260*** *(0.1130)*	-0.0911 *(0.1180)*
DMS × POST (β_3)	12.354***	1.690	+0.1080	-0.0128

	(4.311)	(1.707)	*(0.1600)*	*(0.1670)*
Interaction P-value	0.0048	0.3238	0.5000	0.9390

Notes: Standard errors in parentheses. * p < 0.10, ** = p < 0.05, *** = p < 0.01.

The table displays the parameter estimates from the quasi-experimental Difference-in-Differences framework across four distinct dependent variables: raw counts of total accidents (All Accidents), severe crashes (Major Accidents), and their respective logarithmic transformations (Log All and Log Major). Standard errors are reported in parentheses immediately below the corresponding coefficients. The interaction term (DMS times POST) represents the estimated causal impact of the digital message sign installation.

8. Discussion

The central finding of this research is that once baseline scale distortions are properly controlled for, the deployment of the Dynamic Message Sign (DMS) had no statistically significant structural impact on traffic safety. While a naïve linear framework misleadingly implies a severe hazard spike (+12.35 accidents/month), transitioning to a robust log-linear specification reveals a statistically non-significant **11.4% relative increase** in overall accident frequency. Similarly, the causal effect on major injury crashes remains entirely neutral and statistically indistinguishable from zero (β_3= -0.0128, p = 0.9390). These results demonstrate that the infrastructure intervention neither materially degraded nor actively improved localised roadway safety ceilings on the Denny Way corridor.

The stability observed in major accidents involving injuries is consistent with traffic dynamics inherent to dense municipal networks. Severe collisions are often the result of high-speed impacts or significant structural failures in traffic flow, which are structurally less common in congested city centres than on high-speed highway corridors. The entirely neutral log parameter for severe crashes suggests that the deployment of the dynamic message signage architecture did not induce extreme reckless behaviour or systemic neglect of traffic signals.

This dynamic interfaces with behavioural insights from Hall and Madsen (2022), who posited that highly salient roadside visual metrics can capture immediate attentional capacity, shifting processing capacity away from localised driving conditions. The Denny Way corridor is already heavily saturated with baseline visual stimuli from commercial signage, transit signals, and dense multimodal traffic patterns. The introduction of the DMS may have pushed certain operators toward cognitive limits. However, the slight, non-significant upward variance in overall accident frequency warrants close conceptual evaluation within a distraction framework. In a dense arterial like Denny Way, traffic is heavily characterised by frequent stops, pedestrian crossings, and streetcar tracks. In such high-traffic environments, even a minor, localised fluctuation in driver attention—such as looking up to process messages on the sign—can dictate minor, low-speed "fender-benders."

This threshold was further complicated by simultaneous compounding operational shifts

executed across the Denny Way network throughout 2023, including the implementation of "No Turn on Red" restrictions and recalibrated traffic signal intervals. While highway-centric literature in environments like Virginia (Cardazzi, 2021) suggests that active dynamic message displays systematically elevate crash risks compared to a deactivated one, data indicate that any localised attentional friction introduced by the display was ultimately small enough that the model cannot distinguish it from zero.

One possible conceptual explanation, rather than a definitive empirical finding of this study, is that the real-world informational utility of the display (e.g., transit travel times) may have balanced out the minor visual distraction it generated, resulting in the observed neutral impact. The results of this study indicate that the DMS had no statistically significant effect on accidents in Seattle's urban context, either for total collisions or for those involving injuries.

These findings introduce nuanced policy implications for urban planning and the ongoing deployment of Intelligent Transportation Systems (ITS). To mitigate inefficiencies and prevent unnecessary attentional load, municipal transportation agencies must exercise caution when prioritising high-cost electronic displays over static signage on dense urban arterials, where the real-time marginal benefit of travel telemetry is structurally limited. When DMS networks are utilised, messaging protocols should be strictly restricted to basic, single-line formats to minimise glance durations. Furthermore, rather than relying on physical, roadside display infrastructure that adds visual clutter to high-risk intersections, future transit frameworks should focus on pre-trip digital integration, leveraging vehicle telematics and mobile routing applications that allow drivers to internalise corridor data before engaging with active roadway hazards.

A notable data limitation in this study is the absence of continuous, time-varying traffic exposure metrics, such as monthly Vehicle-Miles-Traveled (VMT) or segment-level Average Annual Daily Traffic (AADT), across the 79-month sample timeline. In traffic safety literature, accident counts are ideally normalised by vehicle exposure to isolate behavioural shifts from changes in baseline congestion. Because localised municipal registry data did not track dynamic volume fluctuations across the entire surrounding surface grid, this study relies on the Difference-in-Differences group indicator to structurally capture and absorb these stationary disparities in baseline traffic density between the high-volume control radius and the localised treatment corridor.

Future research can extend the current empirical framework by incorporating a broader array of DMS deployment locations across Seattle and other major metropolitan areas. Expanding the analytical panel will determine if this statistically neutral 11.4% relative variance represents a consistent systemic urban trend or if it is a localised anomaly unique to the Denny Way corridor layout. Additionally, future research could benefit significantly from more data regarding the specific content protocols displayed on the electronic sign during specific accident windows.

Isolating whether distinct message formats (e.g., traffic safety slogans versus direct congestion travel-time alerts) induce varying degrees of traffic remains a vital avenue for

optimising municipal transport policy.

It is worth noting a major structural limitation inherent to a strictly linear DiD framework when tracking bounded variables like traffic collisions. Because the control group experienced a sharp, absolute downward trend before 2023, projecting this steep, linear baseline into the post-treatment period forces the counterfactual line to technically cross the zero boundary, predicting an impossible negative number of monthly crashes on Denny Way.

Because real-world collisions are bounded at zero, the actual observed post-treatment data points sit mechanically higher than this artificial negative baseline, creating a severe scale distortion that artificially generates a false-positive hazard spike (+12.35 accidents/month, $p = 0.0048$). Once this boundary issue is corrected using our primary log-linear framework—which successfully handles bounded data and satisfies the parallel trends assumption ($p = 0.7367$)—the apparent driver distraction effect completely evaporates, settling at a statistically non-significant relative variance ($p = 0.5000$).

While the log-linear specification serves as our primary model, alternative functional forms and linear trend checks were estimated to verify the stability of the baseline. These supplementary empirical checks are detailed in Appendix B and Appendix C.

9. Conclusion

This study provides empirical evidence that the installation of a Dynamic Message Sign (DMS) on a high-complexity urban arterial had no statistically significant impact on traffic safety. While the linear specification implied a large positive effect, it violates the parallel trends assumption. When baseline scale asymmetries are strictly controlled for using a mathematically valid log-linear specification, the true causal impact drops to a statistically non-significant **11.4% relative increase** in overall collision rates, while severe injury crashes remain entirely neutral ($p = 0.939$).

As municipal authorities continue to expand Intelligent Transportation Systems (ITS), these findings demonstrate that roadside digital signage architecture neither materially degrades nor actively improves localised roadway safety on dense surface streets. Consequently, these results argue for a careful attention to functional form in difference-in-differences analyses of bounded count data.

Given that these findings are derived from a single corridor, municipal transportation agencies should prioritise comprehensive, site-level evaluations across a broader array of arterial layouts before committing high capital resources to physical dynamic displays. Where safety benefits remain unestablished locally, future urban transit frameworks could tentatively explore low-cost, pre-trip digital integration and streamlined traffic flow optimisation to enhance network efficiency without introducing unnecessary visual clutter to high-risk municipal intersections.

10. Appendix A: Data Dictionary

Variable Name	Type	Description	Source
All Accidents	Dependent	Total monthly count of traffic collisions.	SDOT
Major Accidents	Dependent	Monthly count of collisions resulting in at least 1 injury.	Filtered
Log All	Dependent	Total monthly count of traffic collisions in the form of Ln (1+ all accidents)	Calculated
Log Major	Dependent	Monthly count of collisions resulting in at least 1 injury in the form of Ln (1+ major accidents)	Calculated
DMS	Independent	Group Indicator (1 = Denny Way, 0 = 1-Mile Radius).	User Defined
POST	Independent	Time Indicator (1 = After Feb 2023, 0 = Before).	Seattle Geo Data
DMS × POST	Interaction	The DiD estimator captures causality.	Calculated
DMS × MONTH ID	Interaction	The DiD estimator for the Parallel Trends Assumption and the Robustness Check	Calculated

11. Appendix B: Full-Timeline Difference-in-Differences Regression Outputs

Variable	All Accidents	Major Accidents	Log All	Log Major
Intercept (β_0)	47.327*** (1.859)	16.469*** (0.736)	3.7850*** *(0.0690)*	2.7628*** *(0.0720)*
DMS (β_1)	-39.837*** (2.629)	-14.000*** (1.041)	-1.7577*** *(0.0970)*	-1.6284*** *(0.1020)*
POST (β_2)	-13.982*** (3.048)	-1.849 (1.207)	-0.3260*** *(0.1130)*	-0.0911 *(0.1180)*
DMS × POST (β_3)	12.354*** (4.311)	1.690 (1.707)	+0.1080 *(0.1600)*	-0.0128 *(0.1670)*
Interaction P-value	0.0048	0.3238	0.5000	0.9390
Significance	Significant	Non-Significant	Non-Significant	Non-Significant
Observations (N)	156	156	156	156
R Squared (R^2)	0.6692	0.6356	0.7693	0.7309

Notes: Standard errors in parentheses. * p < 0.10, ** = p < 0.05, *** = p < 0.01

12. Appendix C: Pre-Treatment Parallel Trends Analysis

Variable	All Accidents	Major Accidents	Log All	Log Major
Intercept (β0)	64.190*** (3.894)	19.610*** (1.585)	4.0466*** (0.1318)	2.8993*** (0.1406)
DMS (β1)	-53.600*** (5.507)	-16.970*** (2.242)	-1.7030*** (0.1863)	-1.7237*** (0.1989)
MONTH ID (β2)	-0.6746*** (0.1356)	-0.1255 (0.0552)	-0.0105** (0.0046)	-0.0055 (0.0049)
DMS × MONTH ID (β3)	0.5504*** (0.1917)	0.1187 (0.078)	-0.0022 (0.0065)	+0.0038 (0.0069)
Interaction P-value	0.0051	0.1317	0.7367	0.5834
Significance	VIOLATED (p < 0.05)	SATISFIED (p > 0.05)	SATISFIED (p > 0.05)	SATISFIED (p > 0.05)
Observations (N)	98	98	98	98
R Squared (R^2)	0.7201	0.6385	0.8016	0.7472

Notes: Standard errors in parentheses. * $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$

Works Cited

Wardman, M., Bonsall, P. W., & Shires, J. D. (1996). Stated Preference Analysis of Driver Route Choice Reaction To Variable Message Sign Information. https://www.researchgate.net/publication/33038906_Stated_Preference_Analysis_of_Driver_Ro ute_Choice_Reaction_To_Variable_Message_Sign_Information

Hall, J. D., & Madsen, J. M. (2022). Can behavioural interventions be too salient? Evidence from traffic safety messages. https://www.science.org/doi/10.1126/science.abm3427

Ma, J., Smith, B. L., & Fontaine, M. D. (2015). Comparison of In-Vehicle Auditory Public Traffic Information With Roadside Dynamic Message Signs https://doi.org/10.1080/15472450.2015.1062729

Sisiopiku, V. P., Islam, M. M., Wittig, S., & Stavrinos, D. (2014). Perceived and real impacts of digital advertising billboards on driving performance. https://nctspm.gatech.edu/pi/digital-advertising-billboards-and-driver-distraction.html

Shbeeb, L. (2023). Assessing the Influence of Roadside Advertising on Traffic Distraction and Its Interplay with Inside Vehicle Distractions: A Cross-Cultural Study in Jordan and Kuwait.

Journal of Transportation Technologies. https://doi.org/10.4236/jtts.2023.134038

Dudek, Conrad L.; Ullman, Gerald L (2006). Dynamic Message Sign Message Design and Display Manual https://rosap.ntl.bts.gov/view/dot/84769

Seattle Department of Transportation. (2025). SDOT Collisions All Years. City of Seattle Open Data Portal. https://data-seattlecitygis.opendata.arcgis.com/datasets/90a68d4709b54327a6bc1dfa1b900f8d_0/ explore?location=47.614571%2C-122.333025%2C11

Seattle Geo Data, Dynamic Message Signs Dataset https://data-seattlecitygis.opendata.arcgis.com/datasets/SeattleCityGIS::dynamic-message-signs- 1/explore?location=47.618463%2C-122.350057%2C18

Cardazzi, Alexander (2021), Messaging and Driving: Unintended Consequences of Dynamic Message Signs in Virginia https://alexcardazzi.github.io/cardazzi_jmp_dms.pdf

Explaining Differences in Human Cognition: A Biological, Cognitive, and Visual Framework By Akshita Tiwari

1. Abstract

Human cognition and thought processes vary significantly across individuals and cannot be explained by a single factor. Therefore, this paper proposes an 8-factor integrated framework to explain the differences in cognitive thinking among individuals. This framework includes biological, cognitive, emotional, and social factors, which interact to shape an individual's perception and decision-making in various situations. Understanding these interactions gives us a more understandable view of human cognition.

This framework is further illustrated through conceptual diagrams integrated throughout the paper.

2. Introduction

Individuals facing the same situations always interpret and respond in their own ways, differently. For example, you and your friend could be stuck in an escape room when one of you starts panicking and the other stays calm and tries to think clearly. Why are your cognitive thinking and reasoning different even when the situation or stimuli are the same? These differences arise from multiple interacting factors. But as former research reveals, they mainly depend on a person's cognitive processes, personality, bias, reasoning and decision-making pattern.

Several brain regions, like the prefrontal cortex, parietal lobes, etc., play a key role in cognitive processing and decision making. This paper will be discussing and explaining an 8-factor framework and how these biases affect our decision making. This integrated framework includes

(i) brain wiring and neural pathways
(ii) genetics and brain chemistry
(iii) cognitive biases
(iv) culture, language and upbringing
(v) personality traits
(vi) memory and interpretation
(vii) emotional state
(viii) social influence.

3. Literature Review

Research suggests that differences in human thinking are influenced by variations in cognitive processes and reasoning styles. Sloman (2005) suggests that individuals rely on internal models of cause and effect rather than formal logic. Research regarding individual cognitive biases indicates that people differ in how they make and process judgments. For

instance, Berthet (2021) examines measurement reliability and argues that traditional cognitive bias tests are unreliable because they rely on a single question which can be misunderstood. On the other hand, Alós Ferrer et al. (2016) investigate the thinking process by analysing response times and whether individuals face conflict between intuition and reasoning.

While these studies highlight how thinking processes are measured, the underlying cause of these variations often comes from the personality traits of the individual. Namely, Aghdaei and Tabrizi provide a comprehensive review of the factors distinguishing humans from one another, while specifically focusing on personality and intelligence. Similarly, Liu et al. (2024) investigated the interplay between personality traits and cognitive distortions by comparing patients with Major Depressive Disorder (MDD) to a healthy group. This experiment demonstrates that neuroticism-related traits are more common within the MDD group rather than the healthy individuals. While the previous research establishes that traits influence our general thinking, Shreeves et al. (2020) extend these ideas to practical problem solving, such as diagnosing faults in complex systems. Additionally, Siritzky et al. (2023) highlight how psychological research methods unintentionally filter out specific subsets of the population.

While personality and research methods explain much of the variation in thinking, it is also important to examine the underlying cognitive architecture that all humans share. Kahneman (2011), Evans (2008) and Tversky and Kahneman (1974) propose that the human brain has two systems- system 1 is for fast, emotional and effortless thinking, whereas system 2 is for slow, logical, and effortful thinking. Apart from the underlying cognitive architecture, Pessoa (2008) and DeYoung (2010) talk about the biological factor, which connects personality directly to the brain's physical structure while the former talks about the connection between our emotions and cognition thinking. Stanovich and West (2000) informs us about how cognitive styles differ in people and argue that differences in reasoning arise because rational thinking is not determined by a high or low intelligence level but also from cognitive styles and biases.

Together, these studies demonstrate that differences in human thinking arise from a combination of cognitive processes, personality traits, and methodological considerations.

4. Methodology

This paper uses secondary research methodology, focusing on the cognitive, biological, and emotional factors of the mind which differ across individuals. To develop an integrated framework, a systematic search was done across academic databases such as Google Scholar, ResearchGate and PubMed.

The selection method prioritized peer-reviewed studies to establish the understanding of core theories and works published within the last five to eight years to provide modern context. Apart from the above, this paper includes references of the dual-system model of cognition given by Daniel Kahneman (2011). Sources were chosen based on their relevance and eligibility. And synthesizing these perspectives, this paper identifies how universal cognitive

models and individual traits interact to shape human thinking.

5. Discussion

Existing research on this topic suggests that differences in cognitive processing cannot be explained by a single factor, but instead emerge from multiple biological, emotional, and cognitive influences. These factors can be understood through an 8- factor integrated framework, created to explain the differences that arise in the cognitive thinking of everyone. This framework includes eight main factors: (i) brain wiring and neural pathways, (ii) genetics and brain chemistry, (iii) cognitive biases, (iv) culture, language, and upbringing, (v) personality traits, (vi) memory and interpretation, (vii) emotional state, and (viii) social influence. Together, these interacting factors shape how individuals interpret information, make decisions, and respond to situations.

This interaction is shown below in Figure 1.

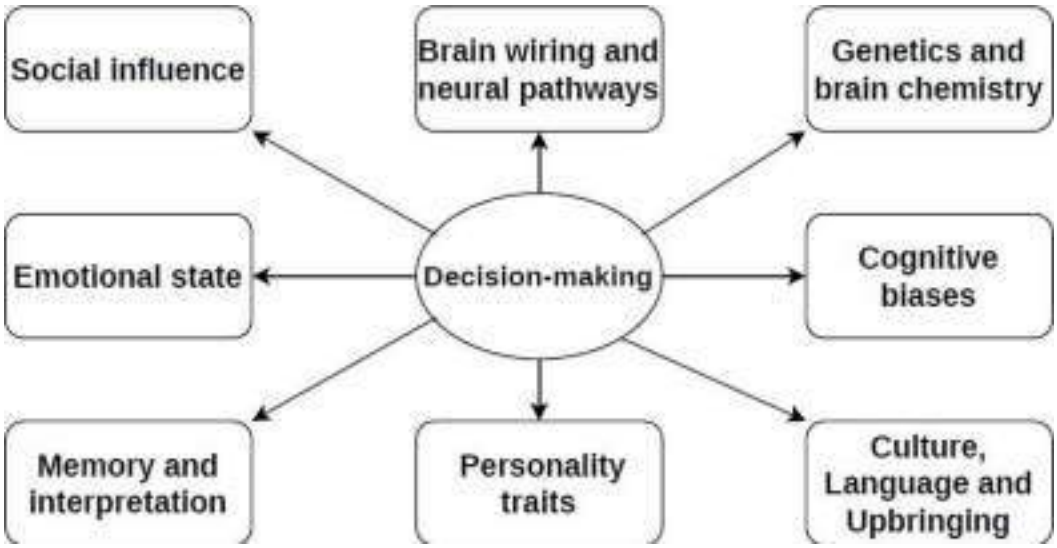


Figure 1: The 8-factor integrated framework which influences an individual's ability of decision-making

(5.1) Brain Wiring and Neural Pathways

According to the principle of Hebbian learning, every time we execute an activity, we activate or fire a network of neurons. When the action is repeated, the pathways become stronger and easier to access; if not, the pathways weaken. This explains the saying "practice makes perfect."

Neural pathways also differ between individuals because key brain regions develop at different rates. For example, the prefrontal cortex, responsible for decision-making and personality, matures differently in everyone based on environment, stress, relationships,

learning styles, and opportunities.

When we examine the brain wiring of different individuals, we also find that each brain forms a distinct default mode network. Neural networks are groups of neurons that coordinate to perform specific functions, such as the

- executive control network – for focus,
- salience network – for observation,
- reward network – for motivation
- and default mode network – for daydreaming.

The default mode network can be understood as the state, the brain returns to when it is not actively engaged in a task, much like the desktop of a computer. It is involved in reflexive thinking and connects regions, such as the medial prefrontal cortex, posterior cingulate cortex, and hippocampus. Together, these regions integrate experiences into a personal narrative. As a result, when two individuals encounter the same situation, each consults a different internal set of memories, beliefs, and predictions, leading to different conclusions.

Another significant factor is synaptic pruning. This is the process by which the brain "deletes" unused neural connections and strengthens frequently used or practiced ones. Synaptic pruning mostly occurs during childhood and adolescence although the timing varies from person to person.

(5.2) Genetics and Brain Chemistry

Genetics provide the blueprint for how our body - specifically our brain - develops and functions. It influences the structure and activity of neural systems along with our wiring and interpretational skills. Genes regulate neurotransmitter systems, neural connectivity, and baseline traits by determining how many hormones are produced, how quickly they are cleared, and how sensitive receptors are. These factors influence mood, reward, motivation, and attention.

Genetics also affect how densely neural pathways are connected, how quickly signals travel, and how adaptable these pathways are. This, in turn, influences cognitive speed, creativity, information processing, and problem-solving style. Additional traits influenced by genetics include emotional reactivity, risk-taking behaviour, and sensitivity to environmental stimuli.

Furthermore, genetics play a role in regulating synaptic pruning, including when it occurs, to what extent, and which connections are strengthened or eliminated. While being mostly guided by genetics, pruning is mostly shaped by experience. Brain chemistry primarily involves the balance of neurotransmitters. The brain contains over 100 neurochemicals that function continuously. These include neurotransmitters, neuropeptides, neuromodulators, hormones that act within the brain, as well as certain gases and amino acids that act as signalling molecules. These chemicals play a deeper role in shaping cognitive processes.

Key neurotransmitters influencing cognition include:

· Dopamine is associated in motivation, reward, and learning from positive outcomes.
· Serotonin is linked to regulates mood, emotional balance, and well-being.
· Norepinephrine (noradrenaline) increases alertness, attention, and stress response.
· Acetylcholine supports attention, learning, and memory formation.
· Glutamate is the main neurotransmitter involved in learning and strengthening neural connections.
· GABA helps regulate brain activity and prevent overstimulation.
· Endorphins relate to reducing pain and creating feelings of comfort or pleasure.
· Oxytocin is released during moments of trust, bonding, and social connection.

(5.3) Cognitive Biases

Human thinking is not entirely logical; instead, the brain often relies on cognitive biases. Cognitive biases are systematic mental shortcuts, often activated unconsciously, that influence how individuals judge, perceive, and interpret situations. These biases allow for faster decision-making by reducing cognitive effort. The study of cognitive biases was developed by Daniel Kahneman and Amos Tversky (1970), whose research demonstrated that human judgment often depends on intuitive heuristics rather than logical reasoning. Even though biases may offer quicker interpretation, these can sometimes lead to errors on judgment. Individuals differ in the types of biases they rely on, influenced by factors such as experience, personality, and problem-solving style. As a result, the same information is understood differently by different individuals.

For example, confirmation bias may lead one individual to favour information that aligns with their beliefs, while another may evaluate the same information more critically. However, cognitive biases are not always detrimental; in many situations, they help in quick and effective decision-making under uncertain situations.

(5.4) Culture, Language, and Upbringing

Culture shapes the norms, values, and belief systems through which individuals interpret the world. It influences how individuals perceive problems, form conclusions, and prioritize goals. Repeated cultural behaviours reinforce specific neural pathways, strengthening patterns of thinking.

Language plays a crucial role in structuring thought. It shapes how experiences, emotions, and relationships are expressed and understood. Language converts sensory experiences such as images, sounds, and feelings into symbols or words. A richer vocabulary allows for more precise distinctions in thought, while a limited vocabulary can reduce the clarity of emotional and conceptual understanding.

Upbringing, particularly during early childhood, strongly influences cognitive development and reasoning style. Factors such as parenting style, education, and exposure to

experiences shape how individuals learn to interpret information and solve problems over time. These experiences can influence the development of brain regions such as the amygdala and prefrontal cortex, which are associated with survival responses and higher-order thinking, respectively. For instance, prolonged exposure to stress may increase cortisol levels, strengthen the amygdala's response while reduce prefrontal regulation.

(5.5) Personality Traits

The consistent pattern of thinking, feeling and behaving that builds up overtime is known as personality, which differs across all individuals. It is affected by genetics, environmental factors, culture and life experiences. Personality shapes attention, perception, behaviour and cognitive style. Stimuli can affect a person's behaviour as shown:

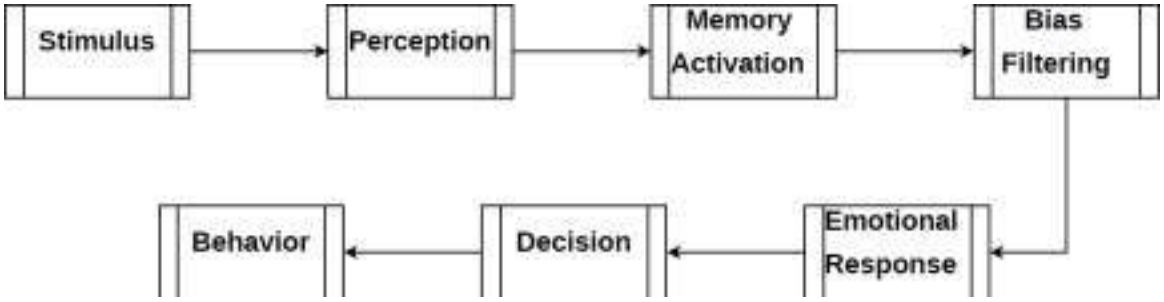


Figure 2: The cause-and-effect relationship between an individual's stimuli which affects their behavior

The Big Five model of personality traits includes:

· Openness to Experience – tendency to notice patterns, metaphors, and novel connections.
· Conscientiousness – main focus on structure and rules.
· Neuroticism – vulnerability to negative outcomes and sensitivity to risk and emotional tone
· Extraversion – leaned towards social interactions
· Agreeableness – sensitivity to emotions and social cues

Individuals also differ in thinking styles, such as:

· Linear (step-by-step reasoning)
· Associative (linking ideas, memories, and insights)
· Analytical (breaking problems into components)
· Intuitive (forming conclusions before explicit reasoning)

Personality influences emotional processing speed. Introspective individuals may experience emotions more intensely and earlier, while pragmatic individuals may prioritize

factual processing before emotional interpretation.

Personality traits affect risk tolerance, impulse control, persistence under stress, and responses to failure. They create default reactions, preferred strategies, and mental shortcuts. However, awareness allows individuals to regulate and override these tendencies. For example, an anxious individual may consciously evaluate their thoughts to maintain objectivity. Personality develops through a feedback loop influenced by:

· Biology (genetics, brain chemistry, sensitivity)
· Early environment (parenting style, emotional safety, validation, or criticism) · Learning and conditioning (reinforcement of behaviours)
· Culture and society (norms, expectations, representations of success) · Experience and micro-level events (e.g., responsibility, failure, recognition) · Cognition
· Choice and self-awareness

The personality loop can be visualised as such:

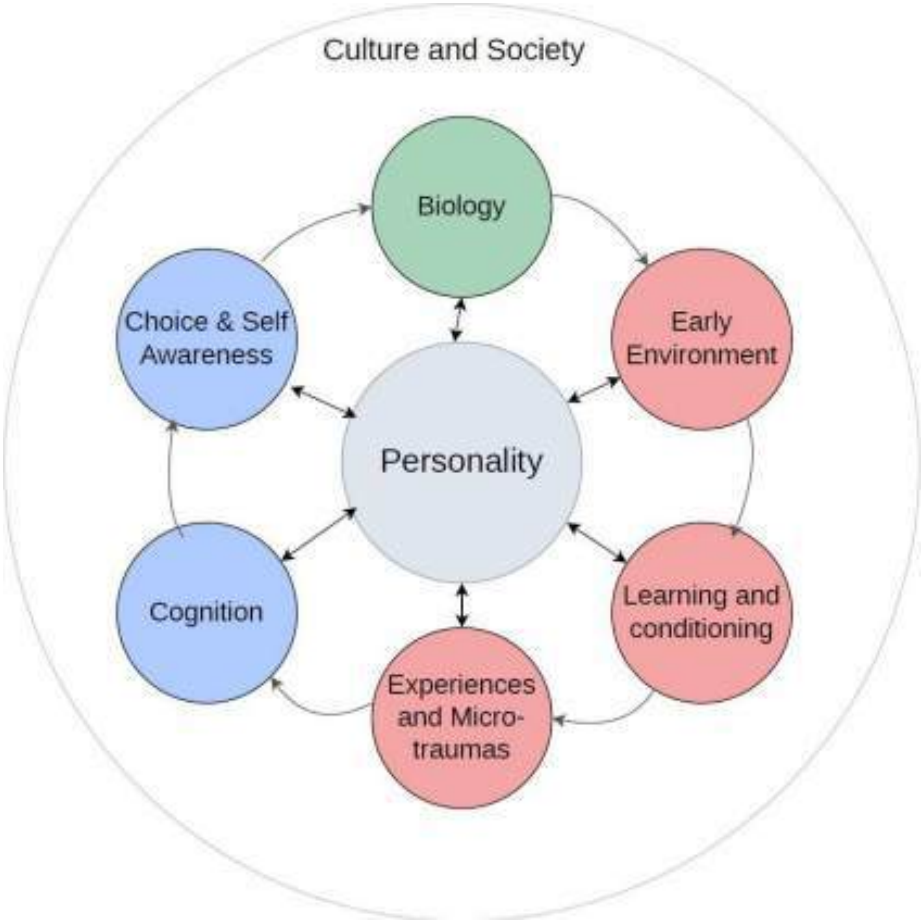


Figure 3: An individual's personality as a dynamic and structured feedback loop.

(5.6) Memory and Interpretation

Memory does not store experiences as exact replicas; instead, it emphasizes emotional intensity, endings, and repeated patterns. When encountering a situation, the brain attempts to identify similar past experiences and retrieves related memories, which may not always be accurate. As a result, a single negative event can outweigh multiple neutral or positive ones.

The brain often predicts rather than passively perceives, sometimes confusing familiar patterns with expected outcomes.

Every memory is associated with an emotional tag:

Strong emotion → stronger recall → greater influence on thought

Interpretation determines the meaning assigned to a memory. While memory provides raw information, interpretation labels and organizes it. An individual's perception shapes these interpretations.

Repeated interpretations can develop into beliefs, which then automatically generate thoughts and influence decision-making. Over time, these beliefs can form narratives. Once a narrative is established, new experiences may be interpreted in ways that reinforce it, potentially limiting openness to alternative perspectives.

This cycle can be represented as:

Thought → Interpretation → Belief → Narrative.

This cycle can be escaped through metacognition, that is being aware of one's own cognition or thought. Instead of being direct reflections of reality, thoughts are predictions made by memory and interpretation. These thoughts ultimately guide decision-making.

(5.7) Emotional State

All our thoughts, perception and our view of the world pass through filters called emotions. They can be categorized into two types:

1. High-arousal emotions (e.g., fear, anxiety, sadness)
2. Low-arousal emotions (e.g., calmness, thoughtfulness)

High-arousal emotional states often lead to faster, more reactive thinking, which can result in less optimal decisions. In contrast, low-arousal states are associated with clearer and more deliberate thinking.

When the brain is heavily engaged in processing strong emotions, fewer cognitive resources are available for analytical thinking.

Emotions often precede reasoning. Individuals may experience a feeling first, and then their thoughts attempt to justify it. This reflects the brain's tendency toward coherence.

Emotions also influence memory. Mood can shape how memories are formed and

recalled, while memories can reinforce emotional states, creating a repeating cycle.

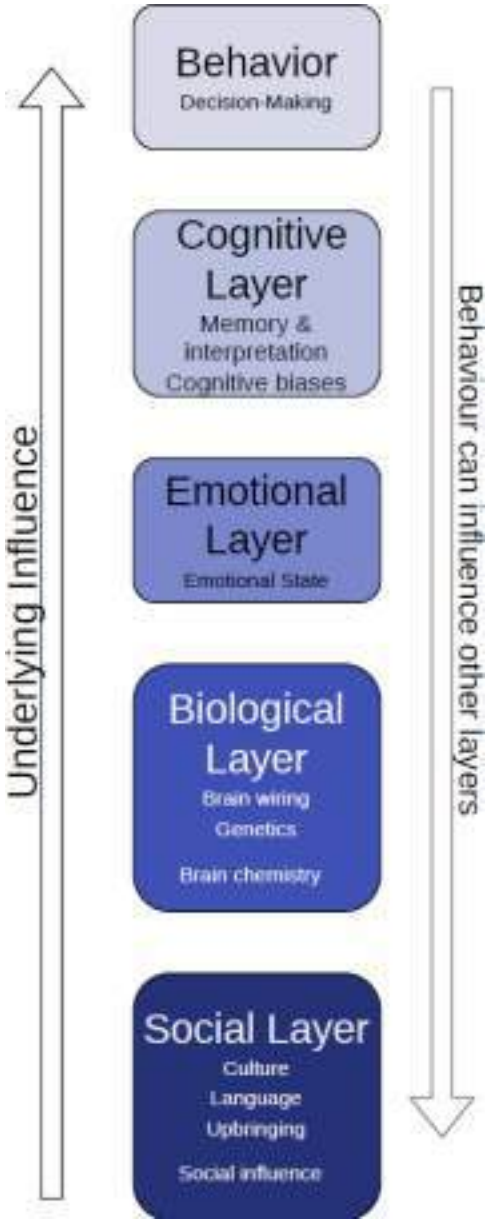


Figure 4: The varying intensity of each layer which impacts our personality and thought process

(5.8) Social Influence

Social influence continues to shape thought processes by establishing norms and expectations that act as default frameworks. These internalized social roles and expectations frequently lead individuals to act based on perceived societal scripts rather than personal preference. While society does not directly control thinking, it establishes boundaries within which thinking occurs. While some decide to operate within these boundaries, some decide to

challenge them.

Language, as discussed earlier, influences thought. When certain emotions or experiences lack linguistic representation, they may become more difficult to identify or express.

Society often promotes comparison, which can distort self-perception, limit ambition, and create pressure to achieve milestones within specific timelines. Individuals may also internalize social roles and expectations, leading them to act based on perceived scripts rather than personal preference. This can reduce independent decision-making.

Additionally, fear of judgment and social expectations can negatively impact cognition and behaviour. While society does not directly control thinking, it establishes boundaries within which thinking occurs. While some decide to operate within these boundaries, some decide to challenge them.

6. Conclusion

Rather than a single cause, cognitive differences arise from the interaction of multiple factors. The 8-factor integrated framework highlights how biological, psychological, and social elements combine to influence cognition, decision-making, and mindset. Because these factors interact in numerous ways, they create countless possible combinations of cognitive processing. As a result, individuals may think in similar ways but not identically, since at least one biological, chemical, or social factor will differ, even if only slightly. Thus, each individual operates through a unique combination of these factors.

Works Cited

Alós-Ferrer, C., Hügel Schäfer, S., & Li, J. (2016). Cognitive reflection, decision biases, and response times. Frontiers in Psychology, 7, 1402. https://doi.org/10.3389/fpsyg.2016.01402

Aghdaei, S. F. A., & Tabrizi, J. S. (2014). A review study of how and why people are different. International Journal of Scientific & Engineering Research, 5(7), 1003– 1006.

Berthet, V. (2021). The measurement of individual differences in cognitive biases: A review and improvement. Frontiers in Psychology, 12, 630177. https://doi.org/10.3389/fpsyg.2021.630177

Liu, Y., Zhang, Y., Chen, X., & Li, H. (2024). Association between personality and cognitive bias in adults with and without depression. BMC Psychology, 12, 58. https://doi.org/10.1186/s40359-024-01634-7

Shreeves, K., Rolls, K., & Walker, S. (2020). Individual differences in strategy use and performance during fault diagnosis. Thinking & Reasoning, 26(3), 330–358.

Siritzky, T., Vankov, I., & Gureckis, T. M. (2023). Standard experimental paradigm designs and data exclusion practices can inadvertently introduce systematic

"shadow" biases. Behaviour Research Methods. https://doi.org/10.3758/s13428-023- 02082-9

Sloman, S. A. (2005). Causal models: How people think about the world and its alternatives. Oxford University Press.

DeYoung, C. G. (2010). Personality neuroscience and the biology of traits. Social and Personality Psychology Compass, 4(12), 1165–1180. https://doi.org/10.1111/j.1751-9004.2010.00327.x

Kahneman, D. (2011). Thinking, fast and slow. New York: Farrar, Straus and Giroux.

Pessoa, L. (2008). On the relationship between emotion and cognition. Nature Reviews Neuroscience, 9(2), 148–158. https://doi.org/10.1038/nrn2317

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. Behavioural and Brain Sciences, 23(5), 645– 665. https://doi.org/10.1017/S0140525X00003435

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. Annual Review of Psychology, 59, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Acceptable Risk, Unacceptable Harm: The Deficiency of Federal Regulation and the Exacerbation of Slow Violence in Cancer Alley By Krish Kocharlakota

Abstract

Industrial pollution in the United States has become a persistent yet often overlooked form of harm, disproportionately affecting marginalized communities through long-term exposure to toxic chemicals. This paper examines Louisiana's "Cancer Alley," an 85-mile industrial corridor along the Mississippi River densely populated with petrochemical facilities and predominantly Black communities, as a key site of environmental racism. It argues that Cancer Alley is best understood through Rob Nixon's concept of slow violence, in which environmental harm unfolds invisibly over time, making it difficult to recognize within conventional legal and media frameworks. Using historical analysis, regulatory case studies, and scientific data on pollutants such as Chloroprene and ethylene oxide, the paper outlines how institutional decisions and regulatory thresholds normalize ongoing exposure as "acceptable risk." The project further explores how media representation and legal processes obscure these harms by privileging immediate, visible forms of violence while sidelining environmental degradation. This paper asks how federal regulation and public discourse render environmental harm a racialized "non-event," and how slow violence provides a framework for understanding the prolonged production of harm in Cancer Alley.

Industrial pollution continues to disproportionately burden many communities in the United States, with growing evidence linking toxic industrial byproducts to severe health issues such as cancer and respiratory illness, as well as long-term environmental damage. Despite advances in scientific understanding of pollutants and the creation of new environmental regulations, the damage persists because industrial pollution in the United States is often minimized as a manageable byproduct, or even a necessary consequence, of economic development. This perspective is reflected in the 1991 World Bank memo signed by economist Lawrence Summers, which stated that "the economic logic behind dumping a load of toxic waste in the lowest wage country is impeccable," revealing how pollution is frequently justified through economic efficiency rather than environmental or human well-being. It is often treated as a problem that can be effectively regulated or even framed as evidence of industrial progress. Yet the actuality is that, for many communities living near sites of heavy industry, exposure to toxic chemicals remains persistent and largely unaddressed, contributing to elevated rates of cancer, respiratory illness, and reduced life expectancy. It is perplexing how regulatory frameworks exist to protect public health, but these conditions still persist and demonstrate how the law is selectively applied. This tension is most apparent in Louisiana's "Cancer Alley," where predominantly Black communities live in close proximity to a high concentration of petrochemical facilities and face some of the highest pollution-related health risks in the nation. Despite the known risks of hazardous substances such as Chloroprene and ethylene oxide, federal regulatory agencies continue to permit their use within so-called "acceptable" limits, effectively normalizing prolonged exposure. In this sense, environmental degradation is not

treated as an urgent crisis but as an ongoing condition. It is rendered as an invisible non-event within the frameworks that fail to recognize its harm. Thus, this paper asks: in what ways does the federal regulation of Chloroprene and ethylene oxide render environmental harm a racialized "non-event," and how does this process exemplify slow violence against marginalized communities?

Louisiana's Cancer Alley is an 85-mile strip of land along the Mississippi River between Baton Rouge and New Orleans. The area contains over 200 petrochemical plants and refineries, making up 25% of the net petrochemical production in the United States. Despite Louisiana being only about one-fifteenth the size of Texas geographically, the state rivals Texas in petrochemical output, concentrating one of the densest clusters of industrial pollutants in the Western Hemisphere into a narrow residential corridor. Even using conservative estimates, Cancer Alley contains roughly 0.2 petrochemical plants per square kilometer, over 25 times the concentration found in the Texas Gulf Coast industrial region, where facilities are spread across a vastly larger geographic area. The EPA reported that the people who lived in Cancer Alley were exposed to over tenfold the level of health risk from hazardous air pollutants compared to the national average. In another investigation, it was noted that only 17% of facilities were in compliance with the Clean Water Act. In a research study from Johns Hopkins University, the primary and most hazardous pollutants found in Cancer Alley were noted to be Ethylene oxide and Chloroprene.

In many ways, life in this region has been sacrificed for the sake of industrialization, making Cancer Alley what many scholars and activists describe as a "sacrifice zone," a residential area significantly damaged by industrial pollution or toxic waste, because rates of cancer caused by air pollution exceed the federal government's own limits of acceptable risk, at rates of 3 times to 10 times higher than the EPA's upper acceptable limit. Furthermore, many have referred to the area as the epitome of environmental racism, where people of marginalized communities are disproportionately exposed to toxins, chemicals, and waste due to race. Studies suggest that African Americans are up to 75% more likely than other populations to live in "sacrifice zones," or communities located near high concentrations of industrial pollution, such as Cancer Alley. For example, "ProPublica reported last year that the EPA does a poor job of regulating the combined risk from multiple sources of industrial air pollution. In parts of Cancer Alley, ProPublica estimated lifetime cancer risk is up to 47 times what the EPA deems acceptable" (Song et al., 2022). And in 1976, when a group of Coast Guard divers in search of sediment samples dove into the Cancer Alley region, they reported back with second-degree burns on their hands, particularly because of the production of vinyl chloride, nitrogen fertilizers, and chlorine in the surrounding area. These "sacrifice zone" framings of Cancer Alley are relevant, and this paper contributes to these framings with a critique of how environmental factors are understood in different United States institutions. The region being considered a sacrifice zone is a blend of abandonment and normalization of degradation. The rates of cancer caused by air pollution alone, exceeding the federal government's own limits of acceptable risk, demonstrate the community's abandonment and normalization.

Cancer Alley may have only received its name in the late 1980s due to petrochemical production. The term simply brings attention to its long-time status as a sacrifice zone, reflecting a repetition of harm that has developed through several transformations of land use along the Mississippi River. The over time changes of the landscape have been inextricably tied to the functions of the Mississippi River. The changes to the landscape, however, were not naturally occurring. The region's proximity to water and its fertile soil have historically attracted different communities. Originally, Indigenous groups such as the Houma, Chitimacha, and Choctaw used the river for fishing, transportation, and other basic necessities for sustenance. However, with the arrival of European colonizers, the Mississippi River quickly became the core of plantation agriculture expansion. This required large amounts of water and fertile soil, which was why the region was crucial. Enslaved Africans were forced to work on these plantations, which ultimately implanted the river into a network of exploitation. Since colonization, this small strip has been defined by resource extraction. Indeed, in the early 20th century, the plantation economy transitioned to an industrial one. The voiding of land for plantations was not simply a single transformation but the beginning of an ongoing process in which the land was required to remain open and controlled. Southern states like Louisiana were able to transform because they leveraged the dispossession of what came before. According to Joy Banner, "Had it not been for the clearing of the land hundreds of years ago for sugarcane plantations, the petrochemical industry today would not have the infrastructure or appropriate land conditions to be sited here." Enslaved Africans labored under conditions that exposed them to environmental hazards that were largely unidentified and unacknowledged. While cancer rates among these populations were never methodically recorded, the absence of data does not translate to the absence of harm. Rather, it reflects a longer history in which their suffering was rendered invisible. Thus, the term Cancer Alley is productive because it announces that harm has always been felt here. Hence, the name "Cancer Alley" captures this reality and underscores the extent of governmental neglect, where communities are effectively left to fend for themselves.

To live in Cancer Alley is to exist under continuous environmental uncertainty that shapes how residents perceive safety and routine. Industrial activity is not an occasional disruption but a constant background condition, requiring individuals to interpret risk through indirect and often imperceptible cues such as air quality changes, odors, or visible emissions. In Cancer Alley, residents report habitual fires at oil refineries. Because the government failed to pre-empt this risk, residents developed their own methods of evacuation in the event of the inevitable next fire. Charles Robicheaux, the occupant of a house at Good Hope, stated that he and his family "had gotten so used to being evacuated because of the frequent fires and explosions at the refinery that they kept packed suitcases next to their beds." Robicheaux's experience reflects how industrial risk is structurally embedded in everyday life, requiring residents to adapt their domestic spaces to anticipate recurring harm. This sustained exposure produces a form of chronic stress in which uncertainty is normalized rather than exceptional, and where concerns about long-term health are inseparable from daily life. Over time, this condition led residents to operate within an environment where harm is accepted but rarely resolved,

reinforcing a persistent imbalance between industrial activity and human well-being. For many residents, prolonged exposure to industrial hazards has undermined the ability to live with long-term security, stability, and environmental safety.

Despite the magnitude of the situation, pollution exposure has been overlooked as a form of violence. In his book, *Slow Violence and the Environmentalism of the Poor,* Rob Nixon (2011) exposed the assumptions tied to notions of violence. Rob Nixon challenges conventional understandings of violence as something "immediate in time, explosive and spectacular in space, and erupting into instant sensational visibility" (p. 2). In other words, violence is flashy, has a perpetrator, and is a singular event. Air pollution, therefore, does not fit neatly within normative understandings of violence. Even though air pollution kills far more than guns. Even if air pollution does not immediately kill, that does not mean that air pollution is harmless. Harm manifests in the diagnosis, the mental health impacts, the reduction in quality of life and life expectancy, as well as death.

The gap between pollution's actual damage and its invisibility became the impetus for a new definition of violence. Slow violence, a term coined by Robert Nixon, refers to the concept of incremental harm that unravels over the course of a long period of time. Slow violence is often imperceptible to the public when it comes to cases of environmental racism or effects that disproportionately impact marginalized communities. This dynamic is deeply racialized because slow violence unfolds gradually and is frequently overlooked by institutions that are structured to recognize only immediate, visible, and attributable forms of harm. The legal system has historically been trained to label Black individuals as perpetrators of violence rather than as victims of it. As a result, the environmental harm experienced by predominantly African-American communities is often not recognized as violence at all but instead reduced to something invisible within frameworks that are not designed to acknowledge this type of inhumanity. Unlike natural disasters, where immediate attention is required, pollution disasters accrue over time through the environments they are forced to live in. Yet the truth is that this notion of imperceptibility is something that has been trained to be ignored; it's not accidental. It is hard to see an elusive company as the perpetrator, especially when the harm can take years to manifest.

On the other side, there is a major bias in how violence is perceived and attributed, particularly in its emphasis on identifiable individual perpetrators. Violent crime, especially gun violence, tends to dominate news cycles and circulate rapidly through media platforms because it fits a simplified and specific narrative structure of a singular offender and an identifiable event. However, this framing is not neutral. An analysis by The Sentencing Project found that crime coverage in television news and newspapers consistently overrepresents Black and Latino individuals as crime suspects while overrepresenting white individuals as victims, reinforcing racialized associations between Blackness and criminality. This distortion is further reflected in newsroom practices, where Black suspects are more likely to be shown in police custody or without individualized context, while white individuals are more often humanized or presented through personal narratives. At the same time, broader criminal justice data reflects deep

structural disparities that reinforce these perceptions. Black Americans are incarcerated at significantly higher rates than white Americans, with Black youth, for example, being more than five times as likely to be detained or committed as their white peers. These combined patterns in media representation and criminal justice outcomes shape public perception, making it more likely that Black individuals are associated with perpetration rather than victimization, even when the underlying harm is structural or environmental rather than individual. This idea sustains racialized logics that repeatedly cast Black individuals as perpetrators of violence, while rendering the slow, structural violence inflicted on Black communities largely invisible.
This bias in understanding violence is also seen with offending chemicals. Hazards such as Chloroprene aren't considered perpetrators or killers because, under criminal frameworks, Chloroprene has only killed one. "To this date, there exists only one single recorded case of acute fatal Chloroprene intoxication, which was a workplace incident in which a worker was found unconscious inside a contaminated vessel and later died from exposure" (Rickert et al., 2011). This lack of direct fatal cases does not indicate safety, but rather reflects the way Chloroprene exerts harm. It is not simply immediate death, but through prolonged exposure that leads to elevated cancer risks over time. Regulatory frameworks use this absence of acute fatalities to treat Chloroprene as manageable, reinforcing the idea that it does not "kill" in any direct sense. In doing so, the state transforms a known carcinogen into an allowed risk, where death is not recognized as the outcome of exposure.

The EPA classifies Chloroprene as a likely human (2B) carcinogen, setting an extremely low recommended exposure level of 0.2 $\mu g/m^3$, showing how it reflects a very high risk of intoxication. Rather than treating this as a motive for the prohibition of the chemical, regulatory practice negotiates exposure, allowing facilities to continue emitting Chloroprene so long as they operate within the threshold limit, although many end up exceeding these limits. In regions such as Louisiana's Cancer Alley, air monitoring has repeatedly shown concentrations far above the EPA's recommended limit, particularly in communities surrounding neoprene production facilities. While scientific assessments show that Chloroprene poses significant health risks at any level, through their inaction, regulatory systems demonstrate that there are acceptable amounts of exposure. This way, Chloroprene does not register as something that directly kills; instead, its effects are structural and dispersed across time through elevated cancer risks and chronic illness. This logic redistributes and normalizes ongoing harm, reinforcing a system in which exposure becomes authorized, and its consequences remain unsung.

Slow violence is lethal, and its threat compounds because pollutants avoid scrutiny. The news cycle obscures the cumulative forms of violence and privileges the spectacle of gun violence. The perpetrators reported in these news cycles are often Black. News cycles prioritize gun violence and cast Black people as the perpetrators of this violence while largely ignoring the slow, structural violence inflicted on Black communities.

Furthermore, these patterns of selective visibility have tangible and material consequences in the continued operation and localized concentration of hazardous industrial production in communities like Cancer Alley. On March 7, 2025, the Justice Department, on

behalf of the Environmental Protection Agency, dismissed a lawsuit against Denka Performance Elastomer LLC concerning its neoprene manufacturing facility in LaPlace, Louisiana, fulfilling President Trump's executive order, "Ending Radical and Wasteful Government DEI Programs and Preferencing." This event underscores how federal action can actively remove efforts to address environmental harm in marginalized communities. Rather than taking accountability for hazardous emissions such as Chloroprene, DEI Programs are considered "Radical" or "Wasteful." This dismissal reflects a broader pattern in which regulatory and legal mechanisms are redirected in ways that allow environmental risk to persist.

The consequences of such regulatory and legal decisions are not abstract; they are directly reflected in the continued operation and localized concentration of hazardous industrial production in communities like Cancer Alley. Chloroprene (2-Chlorobuta-1,3-diene) is a volatile and highly flammable liquid mainly used to manufacture Neoprene rubber, which is commonly found in wetsuits, adhesives, automotive parts, and cable sheathing. The only facility in the United States that engages in the creation of neoprene rubber is located in Cancer Alley. Nearby that facility is an elementary school. The EPA classifies it as a likely human (2B) carcinogen, as it is known to explosively polymerize, catalyzed by the addition of peroxide, when producing the chemical, with the occupational exposure limits generally ranging from 0.55 to 25 ppm. According to the EPA, symptoms reported from acute human exposure to high concentrations of Chloroprene include giddiness, headache, irritability, dizziness, insomnia, fatigue, respiratory irritation, cardiac palpitations, chest pains, nausea, gastrointestinal disorders, dermatitis, temporary hair loss, conjunctivitis, and corneal necrosis. Chronic symptoms as a result of exposure to Chloroprene include the above symptoms, and it has been theorized that it can contribute to liver function abnormalities, disorders of the cardiovascular system, and depression of the immune system.

Ethylene oxide (1,2-Epoxyethane) is a highly flammable and volatile gas used to manufacture items such as antifreeze, plastics, detergents, and adhesives. It is quite often used as a tool for medical sterilization, and is widely known to be a human carcinogen, with the International Agency for Research on Cancer classifying it into a group 1 carcinogen (proven carcinogen). Notably, the direct use of ethylene oxide accounts for only 0.05% of its global production, according to a study done in 2004. This is highly disproportionate when considering its lasting medical effects. Interestingly, the effects on the human body appear to differ by sex, with women showing more noticeable effects than men, specifically regarding breast cancer. In a log-log model with a 20-year exposure lag, workers with cumulative exposure to Ethylene oxide of 3,650 ppm-days experienced a significantly elevated relative risk of breast cancer mortality of 3.15 (95% CI: 1.78–5.60) compared to unexposed workers. Using evidence taken from both humans and animals, it has been shown that exposure to ethylene oxide can result in a large variety of effects. Because it is toxic upon inhalation, it can cause significant damage to the trachea and bronchi, which leads to a partial collapse of the lungs. Higher concentrations of exposure can lead to the buildup of fluid in the lungs, possibly resulting in pulmonary edema or damage to the cardiovascular system. Ethylene oxide can cause acute poisoning, with the

symptoms being headaches, nausea, and vomiting. Chronic exposure has been reported to lead to peripheral neuropathy, impaired hand-eye coordination, and memory loss, even when exposed to amounts as low as 3 ppm. Lastly, ethylene oxide can easily penetrate through commonly worn apparel, causing skin irritation and dermatitis with the development of blisters, fever, and leukocytosis.

Slow violence is, in fact, shaped by the social, political, and economic systems that turn long-term exposure to pollutants in certain communities into a norm, or rather, something with the ability to be dismissed. Instead of being fully unseen, slow violence is actively overlooked. This lack of urgency allows environmental harm to persist without accountability. Agencies such as the Environmental Protection Agency play a central role in maintaining these conditions by defining what levels of exposure are considered "acceptable," even for highly hazardous substances like Chloroprene and ethylene oxide. Rather than preventing harm, regulatory frameworks often normalize it, permitting the continued operation and expansion of facilities that emit carcinogenic pollutants into predominantly Black communities. By doing so, these institutions actively reproduce slow violence, ensuring that long-term exposure remains a condition of everyday life rather than an emergency demanding immediate intervention. Clearly, this persistence is structural and embedded within the institutions tasked with mitigating environmental risk.

In Louisiana, these health risks are not theoretical, as facilities that emit ethylene oxide are concentrated in heavily industrialized regions such as Cancer Alley, where nearby residential communities are directly exposed to these emissions on a continuous basis. In *Inclusive Louisiana v. St. James Parish*, community organizations challenged the local government for approving the construction of industrial facilities in predominantly Black neighborhoods while denying similar development in predominantly white areas over several decades. As the governing authority, St. James Parish played a central role in determining where industrial facilities were sited and, consequently, which communities were subjected to environmental harm. The plaintiffs argued that these decisions produced disproportionate health risks and environmental degradation in Black communities, including sustained exposure to petrochemical pollutants and the destruction of cultural sites such as the burial grounds of enslaved people. The case revealed that, for nearly 50 years, industrial expansion was systematically directed into Black neighborhoods while white communities were shielded, demonstrating that this uneven distribution of risk was not a mere coincidence but structurally produced. However, pollution in this context is not simply a byproduct of petrochemical production; it operates as a colonial force that restructures relationships to land and life itself. As scholars such as Max Liboiron argue, pollution functions as a mechanism of ongoing occupation, limiting how communities can inhabit, use, and sustain their environments. In Cancer Alley, this dynamic forces residents to organize their lives around it through constant adaptation and attempts to repair their surroundings into something livable. This ongoing labor of survival reflects how pollution operates as an instrument of control, ensuring the continued dominance of industrial actors while burdening residents with the responsibility of managing conditions they did not create. Over

time, this labor becomes naturalized, producing a racialized psychology in which the ongoing suffering of Black communities is rendered unremarkable, or something to be endured rather than recognized as violence.

The persistence of slow violence is not incidental but is actively maintained through what can be understood as slow justice: the strategic use of legal delay and institutional process to defer accountability and prolong harm. Companies can afford to sustain long judicial processes yet choose not to in order to serve their own interests. According to Marta Conde, "Slow justice perpetuates and exacerbates slow violence." Judicial processes also pose significant challenges by transforming complex social dynamics into state-defined representations. In court, individuals and communities are reduced to something that is considered "categorizable," ultimately conforming the identities of local and Indigenous groups to state norms and restructuring their demands. As a result, communal histories are diminished into "generic narratives of suffering." This process allows actors to assert control over the lives of marginalized populations, creating psychological distress while further entrenching environmental conflict. This is also reinforced by the fact that corporations can financially sustain prolonged legal battles, while affected residents often rely on overburdened or pro-bono legal representation. As a result, delay itself becomes a form of power, allowing industrial actors to outlast and exhaust community-led efforts for environmental justice. Recent legal challenges against Formosa Plastics further illustrate how slow justice operates, especially within the relationship between the state and industrial actors. While the lawsuit itself functions as a tool for pursuing justice, it also exposes how the law is inherently limiting, as its procedural frameworks and standards of proof often struggle to fully account for the cumulative and structural nature of environmental harm.

In 2026, environmental groups, including RISE St. James and the Sierra Club, sued the Louisiana Department of Environmental Quality for granting another permit extension to Formosa Plastics for its petrochemical complex in St. James Parish, even though the project failed to meet recent federal air quality standards. However, instead of requiring the company to reapply, the state consistently extended permits, allowing the project to remain intact even as evidence showed that it would significantly increase hazardous pollution in a predominantly African-American region. This pattern demonstrates how systems can be leveraged to delay accountability, allowing corporations to avoid conforming to laws while maintaining access to the land. Moreover, environmental reports on Formosa Plastics show how these facilities cause widespread environmental/health harms, including air pollution, water contamination, and long-term climate degradation. The legal process does not simply fail to prevent harm but actively sustains it. Legal institutions often reinforce structural inequalities through procedural delays, strict evidentiary requirements, and prioritization of industrial interests.

Ultimately, this paper has shown that industrial pollution in places such as Cancer Alley cannot be understood as a series of isolated regulatory failures or environmental accidents, but as the outcome of long-standing structural processes that determine where harm is distributed and how it is recognized. Through the framework of slow violence, environmental degradation emerges as a cumulative and racialized condition that is produced through everyday regulatory

decisions, legal structures, and historical patterns of land use that consistently prioritize industrial development over community health. What appears as "acceptable exposure" within regulatory systems is, in practice, the normalization of sustained harm for specific populations, and at the same time, legal and institutional frameworks that claim to address environmental risk often reproduce it through delay, limitation, and selective recognition of what counts as violence. In this sense, Cancer Alley is not an exception but an intensified expression of broader national logics that determine whose suffering is visible and actionable.

These patterns of environmental neglect and uneven protection are not isolated from broader political structures, but instead reflect the same logics of power that shape contemporary struggles over representation and democratic access. Recent debates over voting rights, district mapping, and election control reveal how structural inequities are reproduced through institutions that determine whose voices carry political weight. In Louisiana and other Southern states, controversies over redistricting and the suspension or restructuring of election processes have been criticized by civil rights groups as efforts that dilute Black voting power and reshape political outcomes under the language of "fairness" or administrative necessity. These practices echo the shared underlying logic of how, just as environmental risk is distributed unevenly through regulatory and industrial systems, political power is also redistributed in ways that limit the influence of marginalized communities. From Cancer Alley to contested electoral maps, the same question persists: whose lives and voices are treated as central to decision-making, and whose are systematically minimized? These developments suggest that slow violence is not only environmental but also political, operating through institutions that quietly determine exposure, representation, and ultimately, whose futures are made most vulnerable.

Works Cited

Alvarez, Camila H. "Structural Racism as an Environmental Justice Issue: A Multilevel Analysis of the State Racism Index and Environmental Health Risk from Air Toxics." Journal of racial and ethnic health disparities vol. 10,1 (2023): 244-258. doi:10.1007/s40615-021-01215-0

"Cancer Alley." Wikipedia, Wikimedia Foundation, en.wikipedia.org/wiki/Cancer_Alley. Accessed 27 May 2026.

"Chloroprene." Wikipedia, Wikimedia Foundation, en.wikipedia.org/wiki/Chloroprene. Accessed 27 May 2026.

"Chloroprene (2-Chloro-1,3-Butadiene)." Environmental Protection Agency, www.epa.gov/sites/default/files/2016-10/documents/chloroprene.pdf. Accessed 27 May 2026.

Christopher W. Tessum et al. , PM2.5 polluters disproportionately and systemically affect people of color in the United States.Sci. Adv.**7**,eabf4491(2021).DOI:10.1126/sciadv.abf4491

Environmental Protection Agency. "EPA and State Partners Announce Major Improvement in Clean Water Act Permit Compliance." United States Environmental Protection Agency, United States Government, 22 Nov. 2022, www.epa.gov/newsreleases/epa-and-state-partners-announce-major-improvement-clean-water-act-permit-compliance. Accessed 27 May 2026.

"Ethylene Oxide." NIH, 10 June 2024, www.cancer.gov/about-cancer/causes-prevention/risk/substances/ethylene-oxide. Accessed 27 May 2026.

"Ethylene Oxide." Wikipedia, Wikimedia Foundation, en.wikipedia.org/wiki/Ethylene_oxide. Accessed 27 May 2026.

"From Plantations to Petrochemicals." Sierra, 19 June 2023. Sierra, www.sierraclub.org/sierra/plantations-petrochemicals-juneteenth. Accessed 27 May 2026.

Girard, Jackson. "The People vs. Cancer Alley." Chicago-Kent Journal of Environmental and Energy Law. Chicago-Kent Journal of Environmental and Energy Law, studentorgs.kentlaw.iit.edu/ckjeel/2024/12/06/the-people-vs-cancer-alley/. Accessed 27 May 2026.

"Good-Bye, Good Hope." Facing South, 1 Mar. 1990, www.facingsouth.org/1990/03/good-bye-good-hope. Accessed 27 May 2026.

Groner, Anya. "Cancer Alley." 64 Parishes, 22 Apr. 2024, 64parishes.org/entry/cancer-alley. Accessed 27 May 2026.

Human Rights Watch. "'We're Dying Here' The Fight for Life in a Louisiana Fossil Fuel Sacrifice Zone." *Human Rights Watch*, 25 Jan. 2024, www.hrw.org/report/2024/01/25/were-dying-here/the-fight-for-life-in-a-louisiana-fossil-fuel-sacrifice-zone?utm_source=chatgpt.com. Accessed 27 May 2026.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Re-evaluation of Some Organic Chemicals, Hydrazine and Hydrogen Peroxide. Lyon (FR): International Agency for Research on Cancer; 1999. (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 71.) Chloroprene. Available from: https://www.ncbi.nlm.nih.gov/books/NBK498760/

James, Wesley et al. "Uneven magnitude of disparities in cancer risks from air toxics." International journal of environmental research and public health vol. 9,12 4365-85. 3 Dec. 2012, doi:10.3390/ijerph9124365

Jones, Rena R et al. "Ethylene oxide emissions and incident breast cancer and non-Hodgkin lymphoma in a US cohort." Journal of the National Cancer Institute vol. 115,4 (2023): 405-412. doi:10.1093/jnci/djad004

Kelly-Reif, Kaitlin et al. "Exposure to Ethylene Oxide and Relative Rates of Female Breast Cancer Mortality: 62 Years of Follow-Up in a Large US Occupational Cohort." Environmental health perspectives vol. 133,5 (2025): 57013. doi:10.1289/EHP15566

"Louisiana's 'Cancer Alley' Is More Deadly Than Previously Imagined." Johns Hopkins Bloomberg School of Public Health, Johns Hopkins, 4 Aug. 2025, publichealth.jhu.edu/2025/the-shocking-hazards-of-louisianas-cancer-alley. Accessed 9 Mar. 2026.

Maraniss, David, and Michael Weisskopf. "The Faces of Pollution : As Cancer, Miscarriages Mount, Louisiana Wonders If It Is a 'National Sacrifice Zone.'" Los Angeles Times, 24 Jan. 1988, www.latimes.com/archives/la-xpm-1988-01-24-mn-37913-story.html. Accessed 28 May 2026.

Mikati, Ihab et al. "Disparities in Distribution of Particulate Matter Emission Sources by Race and Poverty Status." American journal of public health vol. 108,4 (2018): 480-485. doi:10.2105/AJPH.2017.304297

Mokhiber, Russell, and Robert Weissman. "Memo Misfire: World Bank 'Spoof' Memo on Toxic Waste Holds More Irony Than Laughs." Global Policy Forum, May 1999, archive.globalpolicy.org/component/content/article/209-bwi-wto/43247.html. Accessed 27 May 2026.

Nixon, Rob. Slow Violence and the Environmentalism of the Poor. E-book ed.

Ostfeld, Jackie. "Your Next-Door Neighbor Should be a Park, not a Smokestack." Sierra Club, 4 Aug. 2023, www.sierraclub.org/articles/2023/08/your-next-door-neighbor-should-be-park-not-smokestack. Accessed 27 May 2026.

Rickert, Annette et al. "A fatal intoxication by chloroprene." Forensic science international vol. 215,1-3 (2012): 110-3. doi:10.1016/j.forsciint.2011.03.029

Smith, Sadie et al. "Social vulnerability and cancer risk from air toxins in Louisiana: a spatial analysis of environmental health disparities." Frontiers in public health vol. 13 1601868. 20 May. 2025, doi:10.3389/fpubh.2025.1601868

Song, Lisa, and Lylla Younes. "EPA Calls Out Environmental Racism in Louisiana's Cancer Alley." ProPublica, Grist, 19 Oct. 2022, www.propublica.org/article/cancer-alley-louisiana-epa-environmental-racism. Accessed 27 May 2026.

Tookmanian, Elise. "Living Near Ethylene Oxide Emissions Linked to Higher Risk of Breast Cancer in Situ." National Cancer Institute Division of Cancer Epidemiology and Genetics, NIH, 25 Apr. 2023, dceg.cancer.gov/news-events/news/2023/ethylene-oxide-jones. Accessed 27 May 2026.

Toxicological Profile for Ethylene Oxide. Atlanta (GA): Agency for Toxic Substances and Disease Registry (US); 2022 Aug. CHAPTER 2, HEALTH EFFECTS. Available from: https://www.ncbi.nlm.nih.gov/books/NBK589511/

United States, Fifth Circuit Court (5th Cir.). Inclusive Louisiana, by and through their members; Mount Triumph Baptist Church, by and through their members; RISE St. James, by and through their members v. St. James Parish; St. James Parish Council; St. James Parish Planning Commission. United States Reports, vol. N/A, 9 Apr. 2025. United States Court of Appeals for the Fifth Circuit, www.ca5.uscourts.gov/opinions/pub/23/23-30908-CV0.pdf. Accessed 27 May 2026.

Weiss, Caroline. "Environmental Injustice: Roots, Impacts, and UrgentSolutions." University of Pittsburgh Dietrich School of Arts & Sciences, 14 Sept. 2023, www.climatecenter.pitt.edu/news/environmental-injustice-roots-impacts-and-urgent-solutions. Accessed 28 May 2026.

Wood, Adrian, and Marilyn Howarth. "How Federal and State Regulatory Systems Perpetuate Environmental Injustice in the United States: Industrial Ethylene Oxide Emissions as a Case Study." Environmental justice (Print) vol. 16,4 (2023): 297-308. doi:10.1089/env.2021.0120

How to Score? Data Analytics Based Strategies for Varsity Level Volleyball

By Koken Sugita

Abstract

This paper collects and uses 27 variables from the UWCSEA boys varsity volleyball team -an international school in Singapore- to examine how data variables can explain the score differential of a game. I defined 17 performance variables across serve, attack, reception, blocking, digs, and errors, and analysed their relationship with set score difference using linear regression in R. The final model suggested that about 88% of the variation in score difference is identified by serve related metrics (effective serves, serve-ins, and service aces) as the strongest positive predictors of success, whilst team faults and errors such as unsuccessful blocks have the largest negative impact. These findings suggest that in a competitive development level, minimizing basic errors and getting consistent but pressuring serves are more important for winning than raw attacking volume, and reflects on how coaches should focus on building fundamentals, serve consistency and block discipline in their players.

1. Introduction

Volleyball, being a fast paced three dimensional sport, requires a blend of athleticism, coordination and strategy. Here, understanding what really motivates success becomes crucial at varsity level, where tiny margins decide who wins and who loses. This is true and relevant for both players and coaches looking to establish a winning system.

Varsity volleyball is a field where small improvements in techniques can make a big difference in outcomes. This paper will propose my findings of what are the important features of a team that if improved can lead to better outcomes and trends associated with those.

To obtain concrete understanding that can be translated into training strategies I need to go beyond speculation and discover the true statistical relationships between performance metrics and winning results. Thus, I gathered actual in-game data from south east asian international varsity boys matches and used R to analyze it. My objective is simple- to give coaches and players who compete in Singapore male Varsity levels a data backed path to improvement and help future teams escape the cycle of losses.

My research question is to discover the key performance variables for varsity-level male volleyball teams in Southeast Asia. In particular, how to define the variables and enable data collection from these variables. In this paper I propose a manual technique watching videos as well as an AI based tool. The analysis and results give insights that can be implemented during training.

My focus is on varsity level (under-19) male volleyball teams. While Marcelino analysed elite men's World League 2005 matches and found that attack efficiency, followed by serve then blocking, to be the most important performance indicators, Da Costa et al. showed that, at school-level volleyball, serve efficiency (high ace counts and low errors) is the most important factor, followed by attack efficiency and then reception efficiency, with setting also emerging as

a decisive factor. Our study differs from Marcelino in that their analysis was based on elite men's international matches in the World League 2005, where the offensive terminal actions such as attack and block dominated set outcomes, whereas for our target high school varsity boys in Southeast Asian international school leagues, a competitive but developmental level. It also differs from Da Costa et al. in that their sample was school female volleyball aged 12–14, with their performance indicators analyzed via error, excellence, and continuity rates, while in our case we define the variables of interest differently. In contrast to both, our methodology follows the approach of Drikos, S. et al. -A evaluative analysis of volleyball skills in balanced sets for men and women in Asian competitions-, which applied statistical modeling (logistic regression, random forest, PCA) to volleyball data from Nations League men's matches to rank the significance of skills, showing that offensive variables were far more predictive of outcomes than defensive ones . Similarly, we used regression analysis but tailored it to our context, collecting around 30 match datasets and running a linear regression model. Our findings identified serve effectiveness as the most significant predictor (lowest p-value), followed by serve-in, serve-ace, and block error. Thus, while Marcelino emphasizes attack as the dominant factor in elite men's play and Da Costa et al. emphasizes serve and reception in school-level girls' play, our study on varsity boys in Southeast Asia demonstrates that serve performance variables are the most decisive in determining match outcomes, reflecting contextual differences across competition levels, gender, and analytical definitions of performance metrics.

The next section discusses the data collection and methodology. The results and discussion can be found in Section 3, while Section 4 concludes.

2. Data and Methods

The methodology uses 17 different variables summarized in Table 1 below. Each represents a component of a volleyball game, to measure and take data effectively. These variables have been carefully selected from various volleyball analytic references; to match the level of play and keep the variable's relevancy, I have excluded the variables which were deemed unimportant or irrelevant through some preliminary regressions. Each variable falls into the category of serve, spike, block, dig, reception, and 'other' variables.

2.1 Data

Data were collected by reviewing videos of 27 sets of UWCSEA dover boys varsity team and recording the occurrence of the events that are summarised in the variables in Tables 1-3, where the definitions can be found.

From SV.LEAGUE, I have used an effective indicator system, which indicates the efficiency in which the skill has been performed. Additionally, I have also decided to other variables, including opponent errors and team faults, as they were good outcome predictors. However, this source specifically measured back attacks, which I decided not to include since back attacks were less frequent in our level of play, and would just be a noise if included.

Variable Definitions

Symbol	Name	Description
Serve Variables		
X_{SA}	Serve Ace	A perfect serve that directly results in a point
X_{ES}	Efficient Serve	A serve that disrupts the opponent's serve receive and prevents a quick attack, but doesn't directly score.
X_{SB}	Serve In	A good serve that stays in play, but allows the opponent to run a quick with all attacking options.
X_{SE}	Serve Error	A failed serve (net or out) that immediately gives the opponent a point.
Spike Variables		
X_{AK}	Kill Spike	An attack that leads directly to a point.
X_{AA}	Successful attack	A spike that keeps the play alive but doesn't score or lose a point.
X_{AE}	Unsuccessful attack	A spike that leads to a lost point (e.g., blocked or hit out).
Reception Variables		
X_{RA}	Reception A	Perfect reception allowing the setter to jump set without moving — enables middle attacks.
X_{RB}	Reception B	Good reception where the setter moves but can still run a quick attack.
X_{RC}	Reception C	Poor reception that limits the setter's options (no middle) and makes the attack predictable.
X_{RE}	Reception error	A failed reception that gives point to the opponent team.
Block Variables		

X_{BA}	Successful block	Block shut that scores the point, or a good block touch that keeps the ball in play.
X_{BE}	Unsuccessful block (error)	The blocker doesn't jump to block, creates a gap in the block, or doesn't press enough which leads to the opponent's point.
Dig variables		
X_{DA}	Dig success	A player successfully receives a spike, and play continues.
X_{DE}	Dig unsuccessful (error)	A player is in the correct spot but fails to keep the ball in play.
Other variables		
X_{OE}	Opponent error	Points gained from unforced mistakes by the opponent (not directly caused by our team's action).
X_{TF}	Team faults	Points lost due to an accident or unforced individual mistakes by our team. (ie, double contact, net touch, line fault).
Table 1: Variable Definitions		

For each set I have recorded the score of each team. Then I computed the outcome variable Y_i, $i = 1,\ldots, 27$ as the score difference for the set. A positive Y_i means a win, Y_i means a loss. I have decided to use this variable as a determinant, since it not only describes a win or loss, it also communicates the margin of it.

2.2 Models

In the analysis I used the classical Linear regression model Bobbitt -How to interpret regression output in R. Statology. I fitted an initial model with all variables to assess significance. The results can be seen in Table 2 below:

Significance Rank	Variable	Estimate	p-value	What it means
1	Effective.serve	+1.09	3.48e-06	Strongest impact. Clean serves that pressure the opponent and prevent quick attacks are key to scoring.
2	Serve.in	+0.81	0.0007	Even just getting the serve in helps. Shows how consistency in serve gives you a solid chance to win the point.
3	Other.Errors	-0.96	0.0014	Very significant. Net touches, double contacts, bad sets — all drop the score big time.
4	Service.ace	+1.33	0.0026	Aces give you a massive advantage — over 1.3 points worth on average.
5	Unsuccessful.Block	-1.06	0.0069	A failed block, like not jumping or leaving a gap, costs nearly a full point.
-	(Intercept)	-9.74	0.016	The predicted score when all variables are 0. Just a baseline, doesn't mean much in practice.
7	Reception.error	-0.51	0.169	Slight negative effect, but not significant. Reception errors might impact, but this data doesn't prove it strongly.
8	Successful.Attack	-0.16	0.258	Not significant. More successful attacks (not scoring spikes) didn't link clearly to more points — maybe poor execution or bad setups dragged it down.

Table 2: Initial results using all variables

Subsequently, I removed all variables that were not significant at 20% level. The model was then fitted again with the remaining variables. All models were fitted in R Studio. The final model can be written as:

$$Y_i = \beta_0 + \beta_{RE} X^{RE}{}_i + \beta_{AA} X^{AA}{}_i \ldots\ldots + \beta_{SB} X^{SB}{}_i + \epsilon_i$$

We assume that the errors are i.i.d $\epsilon_i \sim N(0, \sigma^2),\ \sigma^2 > 0.$

3. Results

Table 3 below displays the coefficients and the corresponding p-values of the final regression model.

Variable	**Coefficient Estimate**	**P-Value**
Intercept	-9.7375	0.016233*
Reception Error	-0.5057	0.168911
Successful attack	-0.1592	0.258085
Unsuccessful block	-1.0556	0.006926 **
Team faults	-0.9633	0.001350 **
Service Ace	1.3333	0.002633 **
Effective serve	1.0904	3.48e-06 ***
Serve in	0.8082	0.000747 ***
$R^2 = 0.8837$, adjusted $R^2 = 83.85\%$,		
Table 3: Coefficients of the final model		

The residuals show the spread between the predicted and actual values. The closer they are to 0, the better the model's predictions. Here, the error ranges from about -3.88 to +4.53. For optimal results and assurance, a residual should be checked against variable plots, and also be tested for normality. In this model, the residuals were found to be randomly scattered around zero with no clear trend, and their distribution was approximately normal, indicating that the assumptions of linear regression were reasonably satisfied.

R^2 illustrates how much the variation in the dependent variable (e.g., point difference) is explained by the independent variable (e.g., block errors, kill spike). In this model, R^2=0.8837, which means that 88.4% of the outcome variation is explained by the model, inferring that this model captures most of what affects the outcome; the adjusted R-squared penalizes for the number of variables. At 83.85%, it shows the model is still strong even after that correction.

The F-statistic tests if all variables combined will explain the outcome better than just the mean. A model with a high F value and low p-value signifies that the model is statistically significant. Our model has an f-statistic of 9.54 with 7 and 18 degrees of freedom and a p-value of 3.354 x 10^-7, providing very strong evidence that the overall model is statistically significant.

A p-value tells if a specific statistic has a meaningful effect, or in other words, significant. It is the probability that a test statistic takes a value as extreme or more extreme than the current value given that the null hypothesis is true. A low p-value indicates strong evidence against the null hypothesis ($\beta = 0$), meaning the variable has a statistically significant effect. A high p-value suggests the variable is insignificant and likely represents noise. In our case, we found a variation of p-values per each variable; I have removed the variables with high p-values (insignificant) to remove the noises, and the model was then re-estimated to identify which variables have a statistically significant relationship with winning.

As from the results, I have concluded that the serves and errors are the most significant variables in our current game. Effective serves, aces, and even just getting it in all helps the score, and so it is important for a team to practice an effective service. I have also found out that errors kill momentum: unsuccessful blocks and team faults weigh more than just a point, and should be avoided. Surprisingly, the effects of the attack variables were weaker than expected; Therefore, although the spike efficiency was low in the data, it did not significantly influence the probability of winning.

4. Conclusions

This study aimed to identify the key performance metrics and variables that influence the outcomes of varsity level male volleyball teams in Southeast Asian international school competitions and leagues, using data analytics. The data consisted of 27 sets of gameplay which were collected manually, and analysed using linear regression of R. The results identified that serve related variables (effective serves, serve-ins, and serve aces) had the strongest positive effects on score difference, and were found to be the most significant predictors of success. On the other hand, the error related variables such as unsuccessful blocks and team faults had a

substantial negative impact on outcomes, usually costing more than a point. These findings suggest that in varsity level volleyball, creating pressure through serving and minimising errors are more important than the quantity of attacking or successful attacks on itself. From a coaching perspective, the training should consist of improving serve quality, block discipline, and training the basics to maximise the effect of these significant variables. These results highlight a contextual difference between varsity level league and pro-leagues, where consistency and basics plays a greater role than offensive greatness and efficiency. However, this research has some weaknesses. It has a limited scope, does not account for point-by-point level analysis, and has a relatively small sample size. In the future, I hope to expand the dataset across more teams and seasons, incorporate point-by-point and individual analysis through the use of AI instruments, which also helps with reliability.

Works Cited

Bobbitt, Zach. “How to Interpret Regression Output in R.” Statology, 2020, www.statology.org/how-to-interpret-regression-output-in-r/ (statology.org in Bing).

Bujang, B., A. Haqiyah, and M. N. Mohamed. “Analysis of the Technical Ability of the Indonesian Women’s Volleyball Team Sea Games 2023.” Jurnal Pendidikan Jasmani dan Olahraga, vol. 9, no. 1, 2024, pp. 133–40. doi:10.17509/jpjo.v9i1.68822.

Costa, Y. P. da, et al. “Technical and Tactical Performance Indicators Based on the Outcome of the Set in School Volleyball.” Motricidade, vol. 13, 2017, p. 34. doi:10.6063/motricidade.12935.

Drikos, S., et al. “A Comparative Analysis of Volleyball Skills in Balanced Sets for Men and Women in Asian Competitions.” Journal of Human Sport and Exercise, vol. 20, no. 1, 2025, pp. 180–92. doi:10.55860/j7d91004.

James, Gareth, et al. An Introduction to Statistical Learning: With Applications in R. 2nd ed., Springer, 2021.

Marcelino, Rui. “Effects of Quality of Opposition and Match Status on Technical and Tactical Performances in Elite Volleyball.” ResearchGate, 2009, www.researchgate.net/publication/50598406_Effects_of_quality_of_opposition_and_match_status_on_technical_and_tactical_performances_in_elite_volleyball (researchgate.net in Bing).

SV.LEAGUE. “2024–25 DAIDO LIFE SV.LEAGUE MEN Regular Season Matchweek 1/Game 1: SUNTORY Sunbirds vs OSAKA BLUTEON – Match Result and Statistics.” SV.LEAGUE, 11 Oct. 2024, www.svleague.jp.

SV.LEAGUE. WOMEN – Japan Volleyball League SVL Official Website. n.d., www.svleague.jp/en/form/b/32105 (svleague.jp in Bing).

Yudiana, Y., et al. “Student Performance Analysis in Volleyball Learning: The Use of a Modified Volleyball Information System Application.” International Journal of Human Movement and Sports Sciences, vol. 10, no. 5, 2022, pp. 913–21. doi:10.13189/saj.2022.100507.

-

Old Tennis Shoes: How Steinbeck's Simplicity Has Shaped My Journey as a Reader and Member of a Community By Tyler Temerowski

In the novel Cannery Row, Steinbeck writes, “it is the hour of the pearl – the interval between day and night when time stops and examines itself.” When I first read this line, I was struck, I too took a pause, a step back in this chaotic world to examine my surroundings. I looked up from my story and gazed upon my room, in front of me were the hundreds of books on my white bookshelf staring down at me. How did the books get there? I thought.

I was never a reader. In elementary school words had little meaning to me, they jumped off the page like pigeons flying off a busy street sidewalk. In middle school words continued to fly off the page. I would have trouble imagining the pictures that the authors intended to create through their stories. I very much hated reading. This sentiment continued at the start of high school; I was scared to express my thoughts, thinking I’d be wrong and that my classmates would make fun of me because I “interpreted a book wrong.” It’s fair to say I was lost, isolated, and secluded from the joys of reading.

On a Sunday afternoon in my freshman year, I was walking down my staircase to the kitchen. Upon entering the kitchen I saw my father sitting in the breakfast nook typing away on his computer. Outside the clouds had darkened their visages, the rain gently sprinkled on my back porch, and little droplets of water drooped from the gutters. As I approached him, he quietly finished his last sip of coffee, put the cup down and asked me if I wanted to go to the library. I agreed and figured I had nothing to lose. On our drive there the rain had begun to pick up and the sound of the wipers created this fast, rhythmic sound. Traffic was moving unbearably slow. Entering the library, I walked up to the second floor and wandered into the section with the classics. I felt so out of place. Where was I? I thought. What are all these books, who are all these authors? I continued thinking. The hardcover books had protective wraps that were torn apart. The paperbacks that hid among the hardcovers all had creased spines. The bookshelves were old and archaic. It felt like I was in an old manor’s study, surrounded by a bunch of books on, “how to kill Tyler so he doesn’t get to read, for dummies.” I picked up one of these “for dummy” books, The Catcher in the Rye, and I couldn’t put it down. I’ve never read this long in my life - nearly one-hundred pages in two hours. As I closed my book on the one-hundred and first page, I turned to look outside the expansive window adjacent to my chair. Outside the clouds began to break, revealing the encroaching night sky. I turned to my Dad who was sitting on the couch next to me. He glanced at me indicating that it was time to go. With my book in hand, we walked out to the parking lot. Looking up, the clouds had parted, and the night sky was twinkling above. I was so eager to return home to finish my story.

For someone who wasn’t a reader back in ninth grade, Salinger’s The Catcher in the Rye stood out to me because of its style. It’s written in the first person, from the protagonist Holden’s point of view. Salinger uses simple diction: money is called dough, instead of using immense descriptors he will simply state, “Mrs. A. was smart.” This differs from the Russian novel where

someone like Dostoyevsky would write, “my subordinate is of the highest regard, a university man with a job as a privy councillor.” I made that last quote up by the way, but the point is clear.

I also appreciated this novel because Salinger illustrates how Holden is secluded, distant from the world’s chaos. This is evident when Holden contemplates staying in New York, or returning to his old life where he dropped out of high school. Holden describes this situation as, “not knowing where the hell to go.” That was, in fact, the same thought I had when I entered the classics section of the library. I didn’t know where the hell to go.

There is another scene that stood out to me for its sense of isolation. A scene about ducks in a pond. I don’t like ducks. They’re irritating. Loud. I’ve never met a single person who’s actually said, “I like ducks.”

Well, you know the ducks that swim around the little lake in Central Park? In the springtime and all? Do you happen to know where they go in the wintertime, by any chance? Asks Holden.

Where who goes? The taxicab driver responds.

The ducks. Do you know, by any chance? I mean does somebody come around in a truck or something and take them away, or do they fly away by themselves – go south or something?

How the hell should I know? How the hell should I know a stupid thing like that?

Ducks are cyclical creatures. They swim above water, duck down, and back up again. In the fall they migrate south, and in the spring they flock back to the north. It’s a constant repetition of never ending dullness and boredom. It’s like Holden, dull and empty after the taxicab driver shuts him down, or me, still distant from the joys of reading.

In tenth grade I was at a Barnes & Noble in Lower Manhattan when I purchased my first John Steinbeck piece, The Long Valley. The Long Valley is a collection of short stories that is written in a simple form. Steinbeck describes California’s natural beauty and introduces the characters in such a light-hearted manner. I genuinely felt connected to both the people and settings of the stories. There was finally this reading motor in my mind that started turning after fifteen years. Wow, reading is fun. I thought for the first time ever.

John Steinbeck's sense of simplicity is much like Holden’s ducks: creatures of habit. Aren’t we all though? I think about my routine, how my orderly room is so ordered because every morning, happy or sad, sunny or rainy, I clean it. Things must be taken care of or they continuously decay, that’s what entropy tells us, right? So to prevent our own high entropy, we maintain order by being creatures of habit: six in the morning wake up, seven get in the car, do school things, six at night return home, homework, lights out by eleven. I go to the grocery store and I pick the two ripest oranges, because I hate imperfection, and I keep the clothes I wear simple. I have a rotation of quarter zips, shirts, and pants that I go through on a weekly basis.

You can probably find me wearing a salmon-colored quarter zip once every five days. It's this level of order and routine that keeps Steinbeck in check. He creates funny, loveable, routine-oriented creatures that are isolated from the frantic world in their own bubble, in their own community.

I run on the adrenaline of community, the excitement of seeing people I know, it's my escape, my hour of the pearl. Who needs coffee when you have community? Without community, my life would be meaningless and dull, like Holden's ducks. In school, I've been able to build a community: my peers, my relationship with my teachers, and probably the actual building itself. I guarantee that the building knows my name, considering the amount of time I spend in it everyday. But I've also noticed that Steinbeck illustrates this same idea of a community in many of his novels, specifically Cannery Row. Doc is at the heart of this community, he's the town's philanthropist, therapist, shrink: he does it all. When people have problems they go to him to solve them. People will ask for an extra buck and he never says no, even if he knows they will just exploit it on a pint of Old Tennis Shoes. When people are looking for relationship advice, they go to him; when they need to cry, to laugh, or to listen to music on his phonograph, they go to him. In short, he's the unifying force of the town. Without him, the world would be on fire.

I sort of feel like Doc at times, but this is only as of late in my junior year. Every morning I walk into the library at the same time (because…routine), with a smile on my face and a loud "hiiii." I try to be outgoing and converse with as many people as possible. When there is a math test, my phone blows up with missed calls and text messages with people wanting my help. There was this one time in the library, right before a test, where one person sat down to ask me a question for math, and by the end of the period eight more people had shown up at the same table, helping one another, and I was just the facilitator, the Doc, looking over all of them collaborating.

I never used to be like this though. I used to be shy, timid, I never had "real friends." I never went to any social events: I skipped every concert, musical, and school gathering because of my social anxiety. I hated this though. I was isolated, secluded (again like Holden). I didn't know what it meant to "hang out with friends." In a class where I was uncomfortable, my face would turn beet red because of my nerves. It still does at times, but even if it does, I am a lot more comfortable with the words I say and I let it pass. I think it's because I found my community, my people, that in this wild and chaotic world my school is just this one tiny blob in the grand scheme of things, and every individual in this blob is warm and welcoming. It's this comfort that keeps me grounded.

I also think reading Steinbeck has helped calm my nerves. His beautiful descriptions of the natural world make you take a step back, a pause to think about the world we live in. One of my favorite lines in Cannery Row is, "*orange and speckled and fluted nudibranchs slide gracefully over the rock, their skirts waving like the dresses of Spanish dancers*." It's the fluidity of the words that make this sentence come alive, it ebbs and flows much like a dress waves as a dancer twirls. The line "slides gracefully" strikes out to me as this calming factor, it really

embodies this idea of being comfortable in your own shoes, within your own community, and Steinbeck's work is a testament to this.

Growing up reading always felt taboo. I'd be scared my friends would say, "you're a nerd if you read." But as I've grown older, I've always noticed the kids with books in their hands are the smarter and more intellectually inclined students. They are the ones who stand out the most. Because of this, I surround myself with these intellectually inclined students, specifically the friends I've made in my junior year Literary Studies class. Entering this course I was still a bit shy, and still afraid of "being wrong." But as the year went on, we were able to bounce ideas off of one another, discuss how we really felt without fear of "being wrong," and enjoy each other's company in the process of it all. Literary Studies has become our beautiful, little reading corner. An escape, our hour of the pearl.

By the second semester of junior year I feel that I have truly created my own community. A community with my literary studies cohort, with my teachers, and my math mentees. There is this one moment though that encapsulates all of this. If you told me to describe my junior year in one moment, this is that one moment: during a lunch break in the library, I was sitting with a friend and challenged him to solve a crossword I made for the school's newspaper. As people walked by they approached us and asked, "*what's that,"* and I would respond with, *"a crossword puzzle, would you like one?"* And just like my math tutoring, within five minutes I was swarmed with people around me, all asking for hints to the puzzle. When someone asked for a hint, I'd look up, stare at them, smile, and wait until they would collaborate amongst themselves to figure it out. *Wow, music to my ears, people harmonizing about crossword clues. I thought.* I continued this thought by reflecting how I'm no longer afraid to speak in class, to share my ideas about a book. Because at the end of the day I know I'm not "wrong." I've realized that I can be this outgoing person, where people come up to me to ask for help, to share secrets, and to even solve puzzles. I've become less Holden, and more Doc.

Then the noise of the anxious puzzle solvers rose. I continued to ignore them, turning my head to the bookshelf instead. My eyes fell upon Steinbeck's novel *Cannery Row,* where I glanced at it, and then tilted my *gaze* a few books over to Salinger's *Catcher In the Rye.* I stared at the books for a little while, smiled, and thought, huh, guess that's *how my books got here.*

Democratizing Failure Analysis: A Natural Language Processing Approach to Identifying Technical and Operational Bottlenecks in Early-Stage Micro-SaaS Ventures

By Vaibhav Narayanan

Abstract

Background/Objective

Cloud computing and accessible AI tools have enabled solo founders to launch niche "Micro-SaaS" businesses. Despite this accessibility, most of these ventures fail within their first year. Traditional entrepreneurship research relies heavily on retrospective surveys, which are often skewed by recall and social desirability biases, leaving the real-time struggles of solo founders largely unexamined. This study aims to capture these entrepreneurial friction points as they naturally occur.

Methods

We developed a computational framework combining natural language processing (NLP), large language models (LLMs), and human qualitative validation. We analyzed 545 anonymous discussions from the r/microSaaS Reddit community. A deterministic LLM pipeline categorized these posts into three strategic domains—Marketing Constraints, Technical Bottlenecks, and Operational/Legal Hurdles—and assessed the founder's emotional sentiment. These results were then validated through manual human coding.

Results

Marketing challenges were the most common obstacle (44.95%), followed by technical issues (33.03%) and operational or legal hurdles (22.02%). A Chi-Square Test of Independence revealed a statistically significant relationship between the type of obstacle and the founder's emotional state ($\chi^{2}(4, N = 545) = 34.82$, $p < 0.001$, Cramér's $V = 0.18$).

Conclusions

Founders experience significantly higher psychological distress from customer acquisition failures and sudden administrative blocks than they do from programming bugs. This approach provides a real-time methodology for computational social science and offers practical insights for independent developers.

Keywords Natural Language Processing (NLP), Micro-SaaS, Entrepreneurship, Sentiment Analysis, Resource Constraint Theory.

Introduction

Background and Context

The software industry has undergone a fundamental shift. With the maturation of cloud hosting, modular web components, and powerful AI APIs, a single developer can now build and deploy a global application in a matter of days. This shift has given rise to the "Micro-SaaS" ecosystem:lean, specialized software products designed for targeted business niches without the need for large corporate structures or significant initial capital.

Problem Statement and Rationale

However, writing the code is no longer the primary hurdle to business survival. The vast majority solo software ventures fail or stagnate within their first year. While extensive research documents why large, venture-backed startups fail, there is surprisingly little data on the day-to-day challenges that defeat independent, self-funded digital entrepreneurs. A solo founder faces a uniquely intense balancing act, simultaneously serving as the lead engineer, growth marketer, compliance officer, and customer support representative.

Furthermore, a major limitation in startup literature is its reliance on retrospective, self-reported data. Surveys conducted months after a venture's collapse often suffer from recall bias, smoothing over the specific, granular frustrations and emotional low points experienced during active operations. Anonymized online developer forums like Reddit offer a compelling alternative, providing a space where founders transparently document their business struggles without fear of professional repercussions.

Significance and Purpose

This research bridges the gap between computational text analytics and entrepreneurial behavioral theory. By applying modern data science tools to the discussions of solo software developers, this study establishes an unprompted, real-time approach to capturing authentic entrepreneurial experiences. The findings provide startup incubators, accelerators, and independent builders with an empirical roadmap of actual day-to-day failure points.

Objectives

This study has two primary objectives:

1. Develop a scalable pipeline that blends AI-driven text classification with human qualitative verification to extract reliable insights from unstructured text.
2. Empirically map the specific technical and operational bottlenecks that drive Micro-SaaS stagnation, analyzing how different hurdles affect a founder's psychological state.

Scope and Limitations

The scope of this study is strictly focused on independent, single-person Micro-SaaS operators active in digital bootstrapping communities. It excludes heavily funded corporate startups. A primary limitation is the cross-sectional nature of the data, which prevents drawing definitive causal timelines showing whether a specific operational challenge directly caused a complete business shutdown.

Theoretical Framework

Traditional frameworks often attribute startup failure to insufficient market demand, cash depletion, or team friction[1,2]. However, these studies assume an organizational structure with multiple stakeholders and financial reserves. For Micro-SaaS founders functioning as single-person operations, Resource Constraint Theory provides a more accurate lens[3]. This theory suggests that a venture's capability is bounded by its access to human, financial, and technological resources. When operating solo, a bottleneck in any single discipline can stall the entire enterprise[4]. Because resources are strictly limited, every hour spent debugging code is an hour taken away from user acquisition[5].

Independent developers also frequently encounter severe Information Asymmetry[6]. While they possess a deep technical understanding of their product, they often struggle to communicate its value proposition to a broader market, making user acquisition and product-market fit a primary point of friction[7,8].

Natural language processing (NLP) and large language models (LLMs) enable researchers to extract structured insights from large volumes of unstructured communication[9,10]. While digital forums are widely recognized as rich data sources, the field of entrepreneurship has been slow to leverage these tools to study solo developers.

Methodology Overview

We engineered an automated web scraping pipeline to collect unstructured forum communications and processed them through a deterministic NLP engine powered by a large language model. This pipeline categorized the data by strategic bottlenecks and sentiment profiles. The automated categorizations were then validated by human coders and evaluated using non-parametric statistical tests.

Methods

Research Design

This study utilizes an observational, cross-sectional digital text-mining approach to translate qualitative human text patterns into quantitative, structured categories.

Participants or Sample

The sample consists of 545 organic, anonymized discussion threads generated by software developers and independent operators.

Data Collection

Data was extracted from r/microSaaS, a public digital forum for solo software builders. Using an automated scraping pipeline via the Apify platform, we compiled historical threads spanning several months. The raw pipeline gathered post bodies, titles, unique identifiers, timestamps, comment counts, and upvote metrics. Pre-processing scripts cleaned the text by

removing HTML tags, platform-specific formatting strings, and broken hyperlinks.

Variables and Measurements

Each text record was evaluated based on two primary attributes:

● Strategic Bottleneck: The main source of business friction, categorized into one of three mutually exclusive domains: Marketing Constraints (traffic generation, ad spend, SEO), Technical Bottlenecks (software bugs, infrastructure errors, API dependencies), or Operational & Legal Hurdles (incorporation, compliance, payment processor restrictions).

●Sentiment Profile: The underlying emotional tone of the text, classified as Positive, Neutral, or Negative.

Procedure

The workflow followed a structured sequence: Data extraction -> Text normalization and cleaning -> Batch execution through an automated OpenAI NLP engine -> Human-in-the-loop qualitative verification -> Non-parametric statistical testing.

Table 1: Data Pipeline and Feature Extraction Specifications

Processing Stage	**Extracted Elements & Actions**
Source Platform	r/microSaaS (Reddit community)
Extraction Engine	Apify platform automated scraping pipeline
Raw Metadata Gathered	Post bodies, titles, unique identifiers, timestamps, comment counts, upvote metrics
Pre-Processing Script Actions	Stripped HTML syntax tags, removed platform formatting strings, deleted dead hyperlinks

Data Analysis

To eliminate random variation in the text classification, the language model's temperature hyperparameter was locked at 0[10]. For human validation, a random subset of 50 posts was manually coded using open and axial qualitative techniques[11,12]. Inter-Rater Reliability was measured using Cohen's Kappa[13]. Macro-level structural associations between the categorical variables were calculated using a Chi-Square (χ^2) Test of Independence, with effect size evaluated via Cramér's V .

Table 2: Automated Classification and Human Validation Parameters

Validation Metric / Parameter	Value / Methodology
Total Sample Size (N)	545 discussion threads
LLM Temperature Hyperparameter	0 (Enforces deterministic output)
Human Validation Subset (n)	50 randomized posts
Human Coding Methodology	Open and axial qualitative techniques
Inter-Rater Reliability (Cohen's κ)	0.86

Ethical Considerations

The dataset consists entirely of public, unprompted, anonymized crowdsourced records. No private user information, real names, or trackable identifying strings were extracted or stored, ensuring total confidentiality and adherence to ethical guidelines for digital platform research.

Results

Applying our NLP taxonomy to the dataset ($N=545$) revealed a clear breakdown of the challenges facing solo founders.

Table 3: Macro Distribution and Sentiment Profiles of Venture Bottlenecks

Primary Venture Bottleneck	Post Count (N)	Percentage (%)	Dominant Categorical Sentiment
Marketing Constraints (User acquisition, web traffic generation, ad conversions)	245	44.95%	Negative
Technical Bottlenecks (Code bugs, API infrastructure fees, server scaling)	**180**	**33.03%**	Mixed / Neutral
Operational & Legal Hurdles (Payment gateway bans, state incorporation, tax compliance)	**120**	**22.02%**	Highly Negative

Marketing constraints were the most significant hurdle for independent software operators, accounting for nearly 45% of all analyzed discussions ($N=245$). Technical engineering challenges ranked second ($N=180$), making up roughly one-third of the

conversation. Interestingly, while legal and operational compliance barriers were the least frequent (22.02%), they generated a highly concentrated pocket of severe negative sentiment.

To determine whether a founder's psychological state depends on the specific category of business obstacles they face, we performed a Chi-Square Test of Independence. The statistical calculation rejected the null hypothesis—which posited that sentiment distributions are independent of bottleneck types—yielding a score of $\chi^{2}(4, N = 545) = 34.82$, $p < 0.001$. Calculating Cramér's V produced an effect size of 0.18, indicating a statistically meaningful, moderate association.

Table 4: Non-Parametric Statistical Association Metrics (Obstacle Category vs. Sentiment)

Statistical Measure	**Calculated Value**
Degrees of Freedom (df)	**4**
Sample Size (N)	545
Chi-Square Statistic (χ^{2})	34.82
p-value	< 0.001
Effect Size (Cramér's V)	0.18 (Moderate association)

Table 5: Granular Proportional Sentiment Breakdown Across Bottlenecks

Strategic Bottleneck Category	**Positive Sentiment**	**Neutral Sentiment**	**Negative Sentiment**	**Total Row Count (N)**
Marketing Constraints	10% (24)	15% (37)	75% (184)	245
Technical Bottlenecks	15% (27)	65% (117)	20% (36)	180
Operational & Legal Hurdles	5% (6)	10% (12)	85% (102)	120

Technical obstacles generally yielded objective, documentation-focused discussions with high neutral profiles (65%). Developers tend to approach code bugs as structural, logical puzzles with definitive solutions. Conversely, marketing struggles (such as failing to find product-market fit) and operational disruptions (like sudden algorithmic payment freezes) pose a direct threat to the survival of the business, triggering severe negative emotional responses (75% and 85%, respectively).

Discussion

Restatement of Key Findings

This study demonstrates that marketing constraints are the most frequent bottleneck (44.95%) for early-stage Micro-SaaS ventures. Furthermore, statistical validation confirms a significant relationship between the type of obstacle encountered and the level of emotional distress expressed by the founder.

Implications and Significance

These findings challenge the conventional assumption that building the software is the most difficult part of launching a digital startup. While independent developers are highly capable of building functional applications, they frequently struggle to acquire customers. Code errors are treated as manageable problems. In contrast, marketing deficiencies and compliance roadblocks operate largely outside a developer's immediate control, introducing a level of systemic uncertainty that drives founder distress and subsequent business abandonment[7,8].

Connection to Objectives

Both core research objectives were successfully met. We built and validated a scalable computational pipeline ($k = 0.86$ Inter-Rater Reliability) and empirically mapped how operational and non-technical shocks govern a founder's psychological posture.

Recommendations

Startup incubators, micro-funds, and educational programs aimed at young entrepreneurs should consider pivoting away from purely technical curricula. Resource allocation models should prioritize programmatic growth marketing, early customer validation, and compliance infrastructure to better prepare solo builders for catastrophic non-technical bottlenecks.

Limitations

The data originates from a single source, r/microSaaS, which naturally skews toward English-speaking, technically proficient, and digitally active users. Additionally, language models can occasionally struggle to interpret complex sarcasm or extreme linguistic ambiguity within unstructured forum data[10].

Closing Thought

By utilizing automated computational social science to democratize failure analysis, we can identify the silent, real-time friction points that dismantle independent ventures. This approach turns the isolating experience of entrepreneurial failure into structured, actionable data that can guide the next generation of digital builders.

Acknowledgments

[Omitted for Double-Blind Peer Review]

Works Cited

1. C. Giardino, X. Wang, P. Abrahamsson. Why early-stage software startups fail: A behavioral framework. International Conference on Software Business. V ol. 182, pg. 27-41, 2014, https://doi.org/10.1007/978-3-319-08738-2_3.
2. N. Wasserman. The founder's dilemmas: Anticipating and avoiding the pitfalls that can sink a startup. Princeton University Press. 2017.
3. J. Barney. Firm resources and sustained competitive advantage. Journal of Management. V ol. 17, pg. 99-120, 1991, https://doi.org/10.1177/014920639101700108.
4. A. L. Stinchcombe. Social structure and organizations. Handbook of Organizations. Pg. 142-193, 1965.
5. M. Marmer, B. Herrmann, R. Berman, N. Furr. Startup genome report extra on premature scaling. Startup Genome. V ol. 10, pg. 1-50, 2011.
6. G. A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics. V ol. 84, pg. 488-500, 1970, https://doi.org/10.2307/1879431.
7. E. Ries. The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses. Crown Business. 2011.
8. S. Blank. The four steps to the epiphany: Successful strategies for products that win. John Wiley & Sons. 2020.
9. B. Pang, L. Lee. Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval. V ol. 2, pg. 1-135, 2008, https://doi.org/10.1561/1500000011.
10. OpenAI. Gpt-4o-mini inference engine and deterministic language processing architectures. OpenAI Technical Report. 2024.
11. M. L. McHugh. Interrater reliability: the kappa statistic. Biochemia Medica. V ol. 22, pg. 276-282, 2012, https://doi.org/10.11613/bm.2012.031.12. J. Saldaña. The coding manual for qualitative researchers. Sage Publications. 2021.
13. J. Cohen. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. V ol. 20, pg. 37-46, 1960, https://doi.org/10.1177/001316446002000104.

The Influence of Personality on Trust By Nichapa Khoncham

Abstract

Trust is a fundamental component of human relationships, it influences how individuals respond in situations involving uncertainty and risk. This paper explores how personality traits influence trust in relationships. It reviews research that conceptualises trust as a psychological process, particularly through models such as Rempel et al.'s framework of predictability, dependability, and faith. In addition, it explores research connecting specific traits (those in the Big 5) to levels of trust and cooperation through questionnaires and experiments (such as trust games). The evidence shows that personality has a significant impact on how participants will trust each other; moral character-related traits have the greatest impact on establishing trust. This understanding can be used to explain differences between individuals and different social contexts on how they develop trust.

Introduction

Trust can be found all around us, particularly in interpersonal relationships. Trust is critical to having strong connections with others and affects how secure or safe a person feels. Rather than using objective data, many individuals instead rely on their instinct and their gut feelings about personality characteristics and first impressions to form their trust in others. Trust can also be seen as a process that develops over time rather than something that appears immediately. For example, Rempel et al. explains that trust develops through three main stages: predictability, dependability, and faith. Predictability focuses on whether a partner behaves consistently, dependability focuses on whether a partner can be relied on in important situations, and faith refers to having confidence in a partner's motives and the future of the relationship. This suggests that trust is not just one simple concept, but something that grows as relationships develop.

Trust-building relies heavily on an individual's own personality as it affects how trustworthy someone is perceived to be (Stavrova et al.). This is because people often associate certain traits with being caring, reliable, or dishonest. Agreeableness is especially important because it is linked to being cooperative, warm, and supportive, which are qualities that many people want in a partner. Stavrova et al. found that people were more likely to trust someone described as agreeable, while extraversion did not increase trust significantly, even though extraverted people are often seen as sociable and confident. This shows that trust may depend more on traits linked to morality and kindness rather than traits linked to popularity or social status.

People sometimes misjudge whether a person is trustworthy because they form premature opinions of the other person's character based on a generalised impression of the person's personality rather than their actions (Rempel et al.). Although previous research has developed strong models explaining what trust is and how it develops (Rempel et al.), and newer studies suggest that agreeableness may be one of the strongest predictors of trust decisions (Stavrova et

al.), there is still room to explore which traits matter most and why. This literature review will investigate the different personality traits and how they influence trust in relationships. Therefore, this paper reviews the previous research on personality traits such as agreeableness and extraversion and their influences on trust. From previous research, agreeableness is likely the trait that influences trust the most because it is closely linked to warmth, morality, and prosocial motivation (Mooradian et al.). The paper reviews trust as a psychological process, how personality traits influence trust judgments, and why agreeableness may be particularly important in close relationships.

Trust as a Psychological Process

The foundation of all human relationships and social interaction is trust, which is developed through interaction and experiences. People make decisions about their ability to trust others every day, often in situations where outcomes are uncertain.Because trust requires an individual to have the confidence in another individual, even when there is a possibility of their being untruthful, trust is closely related to the concepts of risk and vulnerability. Therefore, trust is not just a single decision or attitude; it is also a process of psychological development based on interaction and experience (Rempel et al.) One influential model of interpersonal trust was proposed by Rempel et al., who describe trust as a dynamic process that evolves as individuals gather information about a partner's behaviour and intentions. Rempel et al. studied trust within romantic relationships using questionnaire measures. They found that trust develops in stages, beginning with predictability and eventually leading to faith in a partner's intentions. This suggests that trust grows gradually as individuals gain experience with a partner's behaviour

The three stages of developing trust, as proposed by Rempel et al., include Predictability, Dependability, and Faith. The Predictability stage relates to the degree that one can predict how someone else will behave based on previous behaviours to anticipate how their partner is likely to behave in future situations. As relationships evolve, individuals begin to identify their partners' dependability and identify whether or not they will be dependable in most important or uncertain times. The final stage of developing trust is the Faith stage, where partners reach a much greater degree of trust. In this stage, they trust that their partners have positive intentions and will continue to act in a way that strengthens the relationship. In this stage, the trust of others has less to do with certain behaviours and more to do with perceived motives and the character of the other partner (Rempel et al.).

This model highlights that trust involves interpreting another person's intentions and character. When it comes to the way people interpret another person's behavior, they won't have complete information to create an accurate assumption on how someone will behave. Therefore, people often rely on cues to help them make an inference about whether the other person will be cooperative or opportunistic. People tend to base their judgement of another person on characteristics that convey moral character like kindness, honesty, and fairness (Brambilla et al.). Such characteristics provide useful information on whether the other person is likely to express

trustworthy behavior. As a result, personality traits will likely contribute significantly to how individuals evaluate potential partners for trustworthiness

Taken together, the literature suggests that trust is not a single, fixed construct but a multidimensional psychological process that develops through repeated interactions, positive expectations, and judgments about another person's intentions and behaviour. Rather than relying solely on past experiences, trust is shaped by how individuals evaluate a partner's reliability, consistency, and future actions. As these evaluations are influenced by individual perceptions, personality traits may play an important role in determining how trust is formed and maintained. The following section therefore examines how differences in personality influence perceptions of trustworthiness and the development of trust within close relationships.

Personality traits and trust

Research on personality has increasingly been used to explain why individuals differ in their willingness to trust others in social interactions. While trust develops through repeated experiences within relationships, studies suggest that stable personality traits also play an important role in shaping trust tendencies (Mooradian et al.; Oskarsson et al.). Among the personality traits linked to trust, agreeableness, honesty–humility, conscientiousness, and extraversion have received particular attention, although evidence suggests that some traits contribute more strongly than others (Freitag and Bauer; Oskarsson et al.).

Most research examining personality and trust has used the Big Five framework. However, some studies have adopted the HEXACO model because it includes Honesty–Humility, a trait associated with sincerity, fairness, and a reduced willingness to exploit others (Ashton et al. 140). To examine how personality influences trust behaviour, many researchers have used a behavioural task known as the trust game (Naber et al.). In this task, one participant (the trustor) decides how much money to send to another participant (the trustee). The amount sent is multiplied before the trustee decides how much to return to the trustor. The amount sent reflects the trustor's willingness to trust, whereas the amount returned reflects the trustee's trustworthiness (Naber et al.). Using this approach, researchers have shown that individual differences between both the trustor and the trustee can influence levels of trust in the interactions (Naber et al.). For example, while Agreeableness is usually associated with individuals willing to trust others (the trustor), honesty-humility is the primary predictor of behavior that is trustworthy (the trustee) because individuals are generally less likely to take advantage of their social partners for personal gain (Ashton et al. 143). Other traits, such as Conscientiousness also plays a role in trust-building and cooperation. As highly conscientious individuals typically comply with social expectations and rules, making them more reliable partners in trust-based relationships (Freitag and Bauer 650).

However, studies suggest that not all personality traits influence trust equally. When individuals first encounter someone, they are more likely to evaluate that person's moral character through traits such as kindness, fairness, and consideration for others, as these characteristics help determine whether someone is likely to act cooperatively rather than

exploitatively (Brambilla et al.). This perception is supported by evidence from trust game studies, which show that trustworthy behaviour is strongly associated with unconditional kindness that reflects an individual's stable disposition rather than simply their reaction to another person's behaviour (Thielmann & Hilbig). Together, these findings suggest that stable personality traits associated with moral character influence both how trustworthy individuals are perceived to be and how they behave in trust-based interactions. The integration of these findings suggests that personality traits provide a larger framework to assist in understanding how and why individuals have such different degrees of trust. Instead of primarily attributing trust to situational experience, it may also be attributed to the consistent way that individuals think, feel and behave affecting how individuals form trusting relationships. Specifically, the traits that signify warmth, cooperation and moral character are likely to influence decisions about trusting within social relationships.

Agreeableness and extraversion

Personality traits play a large role in determining how likely someone is to trust others or how likely someone is to be perceived as trustworthy within a relationship. Although the experience of developing trust occurs over time, there has been research on the effect of stable individual differences on the development of trust.

Evidence shows that personality traits can influence how a person approaches trust and cooperation (Mooradian et al.) For example, Mooradian et al. found that agreeableness was positively related to interpersonal trust, while neuroticism was negatively related. This suggests that individuals who are more cooperative and emotionally stable may be more likely to develop trusting relationships, whereas individuals who experience higher levels of anxiety or emotional instability may approach trust more cautiously. As a result of their cooperative, empathetic and emotionally stable personalities, people may also have a higher propensity to trust others; In contrast, individuals who are more anxious or emotionally unstable may struggle to develop trust due to higher levels of suspicion or insecurity.

Two traits that are commonly examined in the study of interpersonal relationships are extraversion and agreeableness (Stavrova et al.). Agreeableness is associated with being warm, empathetic, cooperative, and a tendency to prioritise harmonious relationships (Mooradian et al.). These attributes are very clear indicators of being kind and caring towards those around you, traits are essential in order to build trusting relationships when trust involves vulnerability. On the other hand, extraversion is associated with being sociable, assertive, and confident. While these characteristics may increase likability and social popularity, they do not necessarily indicate moral reliability or benevolent intent.

The experimental trust game is a behavioural task, participants were given personality descriptions of potential partners and then decided how much they would trust them in the game. Stavrova et al. conducted an experimental trust game by providing personality descriptions of a partner and asked participants how much money they would trust to transfer their money to the described partner. Participants showed a significantly higher willingness to send money to partners who had a description that highlighted agreeableness than they did for a partner viewed

as having high levels of extraversion. This is critical because it indicates trust decisions are significantly more highly influenced by personal traits that are associated with morality and prosocial behaviour than with personal traits that relate to social confidence. In other words, being outgoing does not automatically make someone seem trustworthy. This idea is also consistent with Rempel et al. explanation of trust, particularly the concept of "faith," which involves having confidence in a partner's motives and future intentions. If trust requires believing that someone has good intentions, then traits that signal warmth and cooperation are likely to be especially influential. Overall, this suggests that while several personality traits may relate to trust, agreeableness appears to play a particularly central role in shaping how trust forms within close relationships.

Further evidence from Oskarsson et al. indicates that social trust appears to have a genetic influence by personality, including extraversion. This finding suggests that trust is not solely determined by situational factors or life experiences, but may also reflect stable dispositional tendencies.While extraversion may contribute to an individual's general tendency to trust others, agreeableness appears to be more strongly associated with perceptions of trustworthiness and cooperative behaviour. Together, the evidence suggests that both traits influence trust, but agreeableness may play the more significant role in determining how trustworthy an individual is perceived to be.

Conclusion

This literature review explored how personality traits influence trust in relationships. Trust was first examined as a psychological process that develops gradually through interaction and experience (Rempel et al.). Rempel et al identified that three different types of trusts are based on predictability, dependability and faith, highlighting that trust involves forming expectations about another person's behaviour and intentions. Understanding trust in this way provides a foundation for examining how individual differences may shape trust decisions and perceptions of trustworthiness.

The review then considered how broader personality traits influence trust and cooperative behaviour. Research suggests that stable individual differences can affect both the tendency to trust others and the likelihood of behaving in a trustworthy way (Mooradian et al.; Oskarsson et al.). Studies using experimental trust games further demonstrate that personality traits can influence behaviour in cooperative situations (Naber et al.; Thielmann and Hilbig, "Trust in Me, Trust in You"). Building on this broader perspective, the final section examined the role of specific traits, particularly agreeableness and extraversion. Experimental findings show that individuals described as agreeable are trusted significantly more than those described as extraverted, likely because agreeableness signals warmth, cooperation, and concern for others (Stavrova et al.).

Overall, the research reviewed in this paper suggests that personality traits play an important role in shaping how trust develops within social relationships (Mooradian et al.; Oskarsson et al.). While several traits may influence social behaviour, traits associated with moral character become particularly important when individuals evaluate the trustworthiness of

others (Brambilla et al.). In particular, agreeableness may be especially influential because it signals empathy, cooperation, and prosocial motivation (Stavrova et al.). Understanding these processes may help explain why individuals differ in their willingness to trust others and how trust develops across different social relationships. Future research should adopt longitudinal approaches to examine whether agreeableness continues to predict trust as relationships develop over time. Although the current literature consistently highlights agreeableness as one of the strongest predictors of trust, evidence also suggests that Honesty–Humility and conscientiousness contribute to trustworthy behaviour. Investigating how these traits interact across different relationship contexts and cultures would provide a more comprehensive understanding of personality's role in trust.

Works cited

Brambilla, Marco, et al. “Changing Impressions: Moral Character Dominates Impression Updating.” Journal of Experimental Social Psychology, vol. 82, May 2019, pp. 64–73. DOI.org (Crossref), https://doi.org/10.1016/j.jesp.2019.01.003.

Mooradian, Todd, et al. “Who Trusts? Personality, Trust and Knowledge Sharing.” Management Learning, vol. 37, no. 4, Dec. 2006, pp. 523–40. DOI.org (Crossref), https://doi.org/10.1177/1350507606073424.

Naber, Andrew M., et al. “The Relative Influence of Trustor and Trustee Individual Differences on Peer Assessments of Trust.” Personality and Individual Differences, vol. 128, July 2018, pp. 62–68. DOI.org (Crossref), https://doi.org/10.1016/j.paid.2018.02.022.

Oskarsson, Sven, et al. “The Genetic Origins of the Relationship between Psychological Traits and Social Trust.” Twin Research and Human Genetics, vol. 15, no. 1, Feb. 2012, pp. 21–33. DOI.org (Crossref), https://doi.org/10.1375/twin.15.1.21.

Rempel, John K., et al. Trust in Close Relationships.

Stavrova, Olga, et al. “The Effects of Partner Extraversion and Agreeableness on Trust.” Personality and Social Psychology Bulletin, vol. 49, no. 7, July 2023, pp. 1028–42. DOI.org (Crossref), https://doi.org/10.1177/01461672221086768.

Thielmann, Isabel, and Benjamin E. Hilbig. “The Traits One Can Trust: Dissecting Reciprocity and Kindness as Determinants of Trustworthy Behavior.” Personality and Social Psychology Bulletin, vol. 41, no. 11, Nov. 2015, pp. 1523–36. DOI.org (Crossref), https://doi.org/10.1177/0146167215600530.

———. “Trust in Me, Trust in You: A Social Projection Account of the Link between Personality, Cooperativeness, and Trustworthiness Expectations.” Journal of Research in Personality, vol. 50, June 2014, pp. 61–65. DOI.org (Crossref), https://doi.org/10.1016/j.jrp.2014.03.006.

Ashton, Michael C., and Kibeom Lee. “Empirical, Theoretical, and Practical Advantages of the HEXACO Model of Personality Structure.” Personality and Social Psychology Review, vol. 11, no. 2, May 2007, pp. 150–166, 10.1177/1088868306294907. Accessed 13 July 2022

Investigation Into How the Social Pressure Arising From Social Media's Content Virality Presents Promotional Benefits and Drawbacks for Indie Musicians

By Alexis Roges

Lit Review

The first sign of social media dates back to 1844 through Morse code, according to Maryville University (Maryville University). However, the first formal social communication site officially launched in 1997 with Six Degrees. Six degrees was revolutionary for its time, offering users the ability to have user profiles, friend lists, and use programmed chat systems. Consequently the introduction of Six Degrees and its unique features would go on to reshape public communication as a whole, expanding the development of social media from its original source and into an integral part of society. Aleš Rozehnal, faculty at the Law of Charles University in Prague, found in his 2021 study on the role of social media in society that "[t]he huge development of digital services has changed the world…" and "[i]t has changed the way we communicate, our access to information, and the way we buy and use services" (Rozehnal). Over time social media has continued to adapt and shape the formulated standards of society's climate. Notably in today's landscape, social media worldwide is now used for numerous reasons that expand from anticipated ambitions, a big differing reason being popularity growth. The possibility for popularity through social media has paved the pathway for boundless personal advances. With the vast personal capabilities of social media, many new opportunities for its users have opened -and continue to open- up. Musical artists are one group in particular that has reaped the benefits that media extends. Media offers a wide variety of possibilities for musical artists, these possibilities include but are not limited to: virality, career establishment, and promotional benefits. However, in addition to the grand amount of positives proposed to musical artists, there are multiple drawbacks. In particular, society presents musical artists with mass amounts of pressuring attributes. Nathaniel T. Belcik, a previous Lakeridge High School student advised under Dr. Karen Hoppes, expands on this idea by the discovery that music censorship documents back to after the civil war in 1865 (Belcik). Historically musical artists have continuously been affected by the public's opinion, their careers rely heavily on the media and how their fans react to their creations. In hindsight living based on others' opinions is a subject often found to be stressful, which is an issue further developed in a study by the Record Union, a music distribution service, that interviewed 489 independent artists. They found that 44% of the artists worried about being evaluated by others and 40% felt pressure to deliver (Record Union), concurring that the career of a musical artist in the media atmosphere can be felt as one of hardship– the impacts of public opinions and societal pressures hinder the career of many artists. For instance, with the aspect of societal pressures in mind, a study by Andi Coulter, a professor at Washington and Lee University, explains that music artists now have to rely on online attention and interactions with fans to gain profit (Coulter). From this analysis it can be concluded that in addition to the pressure of being evaluated and having to deliver, music artists in today's age must gauge their careers to follow the climate of digital fame. Instead of following

personal interests within their careers, musicians have to cater to the interests of others. And as social media continues to endlessly extend its reach, the door for a market based economy has been opened.

The aspect of a market economy embedded within social media plays a large role in the current stance of the music industry. Since online media has shifted to a market based economy, musical artists are forced to maintain online personas. This is represented within a study by Dr. Maria Kaya and her colleagues from QUT Business School, in which they exclaim that "a recent study by Carter (2009) reveals that artists who earned the most returns had an online presence on multiple social online sites and services such as MySpace.com, Facebook, YouTube, Last.fm and Twitter and popular music blogs such as The Hype Machine" (Kaya et al.). As a result of the market based economy, careers of music artists now fall into the hands of trends and algorithms, leading to an even greater pressing problem for music creators. Key representation of this issue is showcased in an interview by BBC with musical artist Eminem, in which he stated "I want this attention for this music, but then I want to be able to go in public and I want to be able to eat a [--] sandwich" (Eminem). Eminem's first hand account highlights the overarching matter that the increase of social media is limiting the awareness of the public towards respecting not just boundaries but the wishes of music artists. The lack of public awareness shown is an area within society that is often overlooked and subsequently has progressively heightened as an issue, hence presenting the major urgency for this study's conducted research. By diving into this area of social media, the understudied side of the media's impact on musical artists can be highlighted. The aim of this study is to answer the pressing question of "How Does the Social Pressure Arising from Social Media's Content Virality Present Promotional Benefits and Drawbacks for Indie Musical Artists?"

In 2017, the platform Tiktok was formally released to the global public. Since TikTok's initial release it has proceeded to develop into today's climate as one of the biggest social media platforms, with The University of Maine reporting that TikTok is the most popular social platform in terms of time spent. (UMaine). While in addition, a study on Gen Z's digital consumption by Gyan Prakash Yadav and Jyotsna Rai, two doctoral associate professors, found that Gen Z is in the highest portion of social media's active users stating that the generation is "a vigorous contributor, high consumer of online content." (Prakash Yadav and Rai). Tiktok presents itself as a major contributor for an artist gaining content attraction – it presents potential for mass consumer gravitation and high active interaction rates. Therefore Tiktok presents itself as a valuable site for this study's content exploration due to Gen Z being its primary base, following the activity of Gen Z will offer high quality results due to the group's high media presence and activity rates with artists on social media.

In connection with Tiktok's impact on the music industry, Instagram also stands out as a platform of influence for musicians. Instagram was released publicly on October 6, 2010, the platform was originally conceptualized by Kevin Systrom and Mike Krieger to act as a location based check-in labeled under Buban (Ita, 2026). Fastwording into present day's Instagram, a 2021 study by Brittani Wert on examining parasocial relationships between fans and musical

artists, found that users can now “interact with content and others by liking posts, commenting on posts, replying to Stories, and using direct messaging (DM) to talk to specific users” (Wert). In today’s age, Instagram has changed the landscape of media into a way that integrates artist-fan connection and extends the boundaries of parasocial relationships in a digital space. Additionally, in a Pew Research Center report by colleagues Gottfried and Eugenie Park, it was found that Instagram proved to be the 3rd most popular social media platform and had a higher age diversity rate compared to Tiktok, with the largest portion of users being 18-29 year olds in 2025. (Gottfried and Park). Analyzing both Tiktok and Instagram’s content virality and addressing the social pressure implications of social media, will aid in answering exactly why and how indie musicians are being affected by the media.

This study seeks to deepen the exploration into social media consumption and cover the continuous lack of research geared towards specifically the promotional benefits and drawbacks of social media in regards to indie musicians. Within this study the group of indie artists was narrowed down as the main focus of the analysis because of the inherent lack of exploration of the group. Under the indie status, musical artists face many obstacles due to social media; their careers as sole producers are often hindered and changed by the consumption of media. By addressing the gap of indie artists specifically with a focus on virality impacts, this study hopes to discover exactly how to remedy the consequences that social media presents and shed light on the struggles of an overlooked popularized public community.

Method

The approach decided for this study was an implementation of a content analysis methodology. Content analysis was chosen for this particular study because it aligned with the plan of analyzing interviews with indie artists. Notably a qualitative analysis study done by Marcin Majka, a Doctoral graduate in Physics, focused on him performing an in-depth analysis on five of the main methods of analysis, one of which was content analysis. Marcin had concluded content analysis as “a foundational method, enabling researchers to sift through textual and visual data to identify recurring words, themes, or concepts. Through content analysis, researchers can track changes in discourse over time, revealing shifts in public opinion, media representation, and cultural norms.” (Majka) As a result, the definition for content analysis aligned ideally with the study’s goal to interpret the effects of media. The anticipated results I had for this study was that the promotional benefits social media would offer indie artists would be viewership, sponsorships, and increased sales. I also believed that the drawbacks from media platforms would be geared more around stress related issues and a decrease in mental health. In order to test this hypothesis, I needed to collect specific content to analyze and thus these were the various steps followed.

Procedure

1. Define virality, indie, promotional benefits, and drawbacks
2. Determine how to narrow which musical artists suit the study

3. Select specific interviews
4. Analyze the content for recurring themes

For the purpose of this study indie musical artists will be characterized under the following definition from a study by Tiia Kentonen, on what specifically drives indie artists motivation and their marketability, “artists that either work independently as a solo artist or in a band and that do not have a deal with a record label” (Ketonen). However, expanding a bit onto this definition I will also be allowing artists that aren't and weren’t at the time of the interview connected to large record labels rather than strictly none. The three main major record labels that are being looked out for are Universal Music Group, Sony Music Entertainment, and Warner Music Group. The only exception is the service Artists Without A Label (AWAL) owned by Sony. AWAL is a service where an artist can sign to career deals with rather than strictly assimilation with a record label. Although it was acquired by Sony in 2021 it acts as a distribution business that allows indie artists to keep 100% control of their music. These larger labels are characterized within Parkes P. Winder’s study on the conflict between recording artists and record labels over the rights to music ownership, as often “too controlling with respect to the music into which they pour their time, talent, and emotions” (Winder). Major record labels install control perimeters over a musical artist's career, which prohibits the independence indie musical artists wish to maintain. Thus indie artists cannot be signed with any label that significantly hinders their liberties. Additionally, it was decided to allow artists that had periods without being signed with a major label because the information remains relevant. Even though they may now be signed to a major label, at the time of being interviewed their stance on social media followed the mindset of someone independent.

In further regards to this study, **promotional benefits** will be described as a gain in number of listeners, fans, followers; sales, engagement rate, concert requests, media mentions, (Chetan 134) collaborations, and appearances. Part of this definition is borrowed from a study by Adrian Chetan on defining success in the music industry. My personal definition for **drawbacks** is any virality factors an artist mentions with negative connotations, for instance discussions of mental health, restricted creativity, public backlash, loss of identity, and or toxicity.

The key step necessary to address in this study’s process was a way to determine a specific way to choose which musical artists would be included. With that in mind it became essential to define exactly what virality meant within this study, therefore I borrowed a definition from a 2011 study on defining the process of virality by Jeff Hemsley, a researcher at Syracuse University. In Hemsley’s study he defined virality as *“...a word-of-mouth diffusion process where a message is actively forwarded from person to person, within and between multiple weakly linked personal networks, and is marked by a period of exponential growth in the number of people who are exposed to the message.*” (Hemsley 13). With this basis for the studies definition of virality, it became necessary to implant perimeters for what would be considered essentially going ‘viral enough’. Thus, it was concluded that this investigation would look at indie artists that released albums between the time frame January 2017 to March 2026. The 2017 to 2026

time frame was chosen to account for songs that resurgenced within recent years. Notably Emily Formica, within her study on music resurgence in media, described the resurgence of certain songs as “a recent phenomenon” in which older songs are seeing a major increase in popularity due to the media. (Formica). Accounting for this recent phenomenon helps address overlooked indie artists that experienced fame recently, rather than at the moment they released their album. In order to find albums that fit the 2017–2027 time frame, Spotify’s indie tab was utilized. The indie tab within Spotify allows for a generally simplistic and secure access to a variety of artists and albums that fall under the indie status. Although, it should be mentioned that Spotify doesn’t follow the same Indie definition as this study, rather it incorporates music that is described to be indie by its status. Henceforth, this variation of describing indie was factored in when selecting an artist and Spotify was strictly used for finding names that related to the indie category.

Next a virality benchmark was established, the benchmark focused on the most streamed song in an album and correlated it with the amount of reposts the song accumulated. The first initial step to the benchmark was to find an album by a creator that fell both under Spotify’s indie tab and that upheld the prior definition for an indie artist. As long as both conditions were met, step two was initiated, the next level was to find the most streamed song within the album. Once finding the most popular piece, the research would move onto Tiktok and instagram. Later on either social platform, the song’s title would be put into the search tab to find the amount of reposts a song gained over time. If the song had massed 100,000 or more reposts on at least one platform, it was concluded that there was a correlation between an artist's streaming virality and social media performance. It should be noted that the song could’ve gained attraction from the artists popularity rather than from a trend. However, the correlation between the two factors shows sufficient significance for there to be a cause and effect relationship.

Equation

$$V = (R_T \geq 100,000) \vee (R_I \geq 100,000)$$

- V: Virality achieved.
- R_T: TikTok Reposts.
- R_I: Instagram Reposts.
- $\vee$: The logical "OR" (meaning if at least one side is true, the artist qualifies).

Repost Example

Limitations of Method

Within this study there remains a deal of limitations; one major restriction being that an artist's song could've gained attraction from their popularity rather than from a viral trend. It is quite common for a song to be popularized because of the previously established fame of a singer and their accumulated fan base. Additionally, it can be assumed that the promotional benefits side may be discussed more frequently than the drawbacks of virality, simply because many artists may prefer to dismiss the discussion of negative connotations. Furthermore, it must be taken into consideration that only one of two specific social platforms may have accumulated 100,000 reposts as a result of algorithm variations. Neil Thurman mentions that reasons for variations in algorithms may be specifically attributed by user behavior and differences in consumer appreciation (Thurman et al.). The age differentiation previously established between Tiktok's users and Instagram's members is an indication of such to be true. When it comes to the discussion of the media, the subject of social media tends to be discussed in a broader nature rather than strictly a specific platform. This forced the study to take deeper investigation in order to develop a conclusive outcome. In the current atmosphere of the music industry a portion of indie artists have become signed with major record labels, however if an interview was conducted before their affiliation it still offers valuable information. If an artist wasn't signed at a certain time period this creates a period where their mindset was geared towards that of an independent creator. Additionally an accumulation of a major label must be noted in order for the data to be analyzed accurately.

Analysis Procedure

Collaboratively two colleagues helped to find and analyze 20 different indie musicians, for privacy reasons their names will be credited under Colleagues 1 and Colleague 2. The set amount was decided as 20 since it was a reasonable number for capturing accurate results. After repeating the method process for 20 artists, the selection of who would be researched further was completed. Since there was now a determined group for the study, I could next switch my gears towards analyzing the Primary Data Source. For individual artists, the interviews sourced were taken from either an in person discussion from Youtube and or an article based questionnaire. Within each specific interview with an artist, it was essential an artist discussed the promotional effects and or drawbacks caused by gaining attraction from social media. Therefore, a codebook

for analyzing the criteria of an interview was created. The codebook itself was broken down into seven different sections as displayed below.

Table Sections

(Artist name, album, year, song name)

- At the time of going viral were they in a major record contract?
- Did the song gain over 100,000 reposts? (evidence)
- What did you input in the search bar? (Ex: #Mylovemineallmine)
- Definitions of promotional benefits and drawbacks
- Link(s) to Interviews Watched

Table

(Artist's Name, album, year, song name)	At the time of releasing the album were they in a major record contract? (Universal Music Group, Sony Music Entertainment, and Warner Music Group.)	Did the song gain over 100,000 reposts? (evidence) Edit: Only 1/2 Social Platforms have to have over 100,000 reposts	What did you input in the search bar? (Ex: #Mylovemineall m ine	Promotional Benefits - "Number of listeners, fans, followers; Sales engagement rate, concert requests, media mentions and.. appearances" (Chetan 134).	Drawbacks - Virality factors an artist mentions with negative connotations; mental health, restricted creativity, public backlash, loss of identity, toxicity	Link(s) to Interviews Watched
Mitski, The Land Is Inhospitable and So Are We, 2023, My Love Mine All Mine	No, she was signed to Dead Oceans when she released her album The Land is Inhospitable and so are we.	Instagram - 1M reels Tiktok- 7M Posts	Both - My Love Mine All Mine	Mitski elaborates that she is thankful for the spread of her songs via Tiktok. (Not useable for data information)	In the interview with Raina Douris, Mitski notes that it takes a lot of energy to maintain a public persona, stating that "Heightening certain aspects of yourself so that you can navigate among all these different people" (it's like being at a party).	https://m.youtube.com/watch?v=4VMh6kGETRw&pp=ygUQbWl0c2tpIGludGVydmlld9IHCQnFCgGHKiGM7w%3D%3D

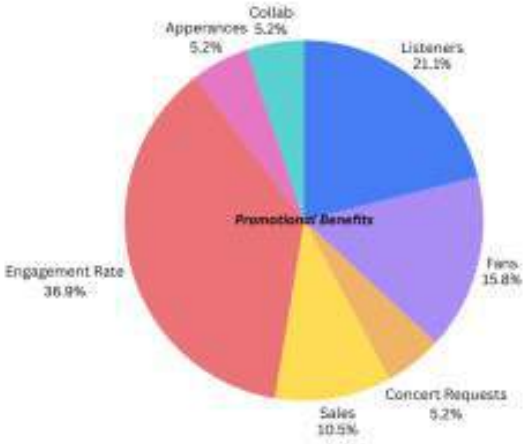
Ethical Considerations

In order to retain accurate representations for each musical artist, ethical measures were adopted. This study relied strictly on interviews available to the public and all quotes were analyzed with their original context to avoid breaching the original context. Furthermore, the anonymity of Colleague 1 and Colleague 2 remained obscure, and every original content creator was properly credited.

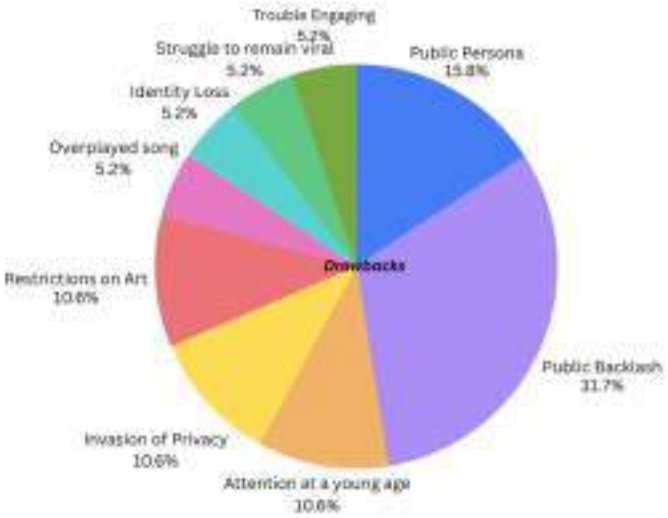
Results

A content analysis approach was used to test the hypothesis that social media's relationship with indie artists pushes the promotional benefits of viewership, sponsorships, and increased sales, as well as drawbacks that are geared around stress related issues and a decrease in mental health. As a result of thoroughly analyzing 20 artists' interviews–reaching a cumulative total of 27 interviews– the results revealed to be a bit nuanced. Each artist provided at least one promotional benefit and or drawback, with a portion mentioning multiple.

Out of 27 interviews, 19 total interviews showcased indie artists responding with answers that signified promotional benefits. Collectively 7 promotional benefits were identified as gained: listeners, fans, concert requests, sales, engagement rates, collabs, and appearances. Statistically the three main benefits that were noted highest were a rise in engagement rates (36.9%), a gain in listeners (21.1%), and an increase in fans (15.8%). With approximately 7 of 19 extracted interviews describing the theme of engagement rates, the artists who mentioned an increase in fan-to-content interactions were Clairo, Cafuné, TV Girl, Chase Atlantic, Djo, Rio Romeo, and Tom Odell. When describing engagement rate, the artists tended to report that people began to resurge and use their music throughout social media. Joe Keery (Djo) in a 2025 interview with Zach Sang commented on his song "End of Beginning" by saying that it "...happened to be at that moment where a bunch of people took it and put it in this Tiktok or put it in this thing" (Keery). In particular, Joe Keery was referencing social media's ability to allow users to put a song in a post and how this spreads notice towards artist's music. Additionally, 4 of 19 interviews featured artists noting an increase in listeners, those of which being Sidney Gish, Samia, Luvcat, and Gigi Perez. Highlighted in a 2025 Stereogum interview, it was cited that Samia had "gotten over a million new streams" within a few days by the time of August 5, 2025 (Jones). A prominently repeated theme in each of these interviews was social media's ability to boost song viewership. Furthermore, a surge in fans was noticed by three artists: Chase Atlantic, Laufey, and Rio Romeo. In a 2024 interview with Ammar Kalia, Laufey stated "We ended up growing a real audience of young people" (Kalia). Laufey was able to grow her audience number with social media's assistance, similar stories appeared with artists Rio Romeo and Chase Atlantic.



In regards to the drawbacks of social media, out of 27 interviews, a total of 19 interviews occurred in which an artist mentioned a negative factor triggered by experiencing mass social media virality. In total 9 different drawbacks were detected: public persona, public backlash, attention at a young age, invasion of privacy, art restrictions, trouble engaging with younger generations, overplaying of an artist's music, struggling to remain viral, and experiencing identity loss. The top drawbacks mentioned were public backlash (31.7%) and a maintenance of a public persona (15.8%). While describing public backlash, the six artists Adrianne Lenker, Gigi Perez, Beabadoobee, Clairo, TV Girl, and Cigarettes After Sex referenced the theme of social media users responding with disapproval or hateful comments. Beabadoobee responded to an interview with i-D that people on the internet can be "quite mean or judgemental" and that "...you're never going to have 100 amazing comments...people on the internet are not real, and ultimately, they do not know who you are" (Beabadoobee). Between the six artists it was prevalent that being present on social media may lead to unanticipated backlash. Additionally, with a rate of 3 out of 19 interviews from Mitski, bbno$, and Sidney Gish, maintaining a public persona was specified. In Mitski's interview with Raina Douris, Mitski noted that it takes a lot of energy to maintain a public persona, stating that it's essentially "[h]eightening certain aspects of yourself so that you can navigate among all these different people" (Mitski). The interviews featuring public personas referred to labels as more of something that was prompted by social media rather than the artists themselves.



Based on the responses recorded from the 27 interviews, the data identified the main promotional benefits as a raised engagement rate, gain in listeners, and an increase of fans, while the top backlashes showcased were public backlash and a maintenance of public persona.

Discussion

In order to answer “How Does the Social Pressure Arising from Social Media’s Content Virality Present Promotional Benefits and Drawbacks for Indie Musical Artists?” a content analysis method was used. The theorized idea was that social media’s content virality would provide indie artists with the promotional benefits of viewership, sponsorships, and increased sales, as well as drawbacks that were geared around stress related issues and a decrease in mental health. Ultimately, the hypothesis was both disproved and validated in certain aspects.

Based on the results, promotional benefits showed that indie artists had the highest gains in engagement rate, increase in fans, and listeners. The results indicate that when an indie artist becomes viral through social media, the rise in public attention creates career opportunities that boost interactions with media users and develop audience members. Sajid Si with his study on social media’s role in marketing, supports this outcome by finding that “[s]ocial media can be useful for creating connections with those who might not otherwise know..” and works to generate exposure and increase traffic/subscribers (Si). This assessment partially refuted my previous claim that social media’s greatest promotional benefits would be viewership, sponsorships, and increased sales. Although the claim acknowledged the gain in listeners partially through viewership, artists primarily talked more about their content interaction. This makes greater sense when taking into consideration that social media is built on user chat systems and content creation.

Furthermore, the drawbacks illustrated that public backlash and maintenance of a public persona were the biggest factors when discussing the negatives of social media’s virality. This outcome suggests that social media can make it difficult for an indie artist to retain their individuality without receiving hate. Thus artists are often forced to find safety and comfort in a persona that the media favors, in order to regulate the public’s disapproval. In support of this evaluation, Waseem Akram and R. Kumar’s study on social media found that “[s]ocial media can without much of a stretch destroy somebody's notoriety just by making a false story and spreading it over the online networking” (Akram and Kumar). When establishing the fact that social media can easily destroy the reputation of an artist, it becomes increasingly sensible that an indie musician would reference drawbacks tailored towards the media’s opinion. Subsequently, because the media can easily damage an artist's career, indie musicians primitively note upholding a likeable appearance and dealing with displeasing feedback. Henceforth, the results disproved the concept that the primary drawbacks would be stress related issues and a mental health decrease by instead focusing on the perception issues artists face.

Implications

Previously established was the implications of a market economy within social media’s atmosphere. The results from this study have the potential to contribute to the overarching strengths and weaknesses of such a system through the understudied section of indie artists. Brittani Wert in her study on parasocial relationships on instagram writes that “...as social media becomes an integral part of modern society and culture, it also permeates the music industry and can become a crucial part of an artist’s marketing strategy” (Wert). It is clear that with the new

found information on the promotional benefits of social media, independent artists can use the system to boost their careers. However, by utilizing social media as an outlet for career success, indie musicians are faced with trials that can prove to be hindrances to the artist and their career progression. The results further imply that public awareness within the media is continuously underdeveloped when it comes to an artist's perspective. Media consumers should understand that artists are personally affected by the consumption and responses towards their creations, in order to foster a greater ethical and sustainable environment for art consumption.

This research was designed to build upon the preexisting marketing economy data by illustrating the specific trade-offs encountered by the understudied section of indie artists. Future research should seek to explore other artists with different statuses of production or specific genres and explore an allocated way that brings higher consumer awareness. These studies should include qualitative and quantitative methods that investigate social media's market economy and the societal pressures pushed by media platforms.

Limitations

This study's results present a variety of limitations that must be taken into account. The primary limitation that must be addressed is the quantity of artists analyzed, this study only considered 20 individual artists. Due to the restricted sample size, the results may not be representative to the entire independent musician population. Thus this study's information should only be used as limited context or brief time period evidence. In addition, this research only addressed the content virality gained from specifically Tiktok and Instagram. Tiktok and Instagram only weigh in as two of many social media platforms. As a result, this limited scope of social media outlets presents platform-specific bias, notably through algorithmic differences and platform structure. There is also a time-bound restriction, the timeline of January 2017 - March 2026 doesn't account for artists before or after the period. Therefore excluding musicians that produced viral music before the period and artists that may gain attraction in the future. While the results don't attribute to every indie artist's experience, this research can still be attributed as context for social media's content virality or brief time period evidence.

Conclusion

Overall, while social media can present ways that boost an artist's visibility, its user base makes it almost impossible for an artist to shine in the public light without criticism and or obstacles. There is a strong potential for media consumers, researchers, and independent artists to use this research to assist in the acknowledgement of how indie artists are affected by social media's quick content virality. Future developments should focus on identifying ways to spread user awareness regarding their impacts through social media and understanding the climate of online media as a whole. As put by researcher Sultănica Crăcană and her colleagues, "[t]here is no escaping social media these days, either for people or for companies." (Crăcană et al.).

Works Cited

Akram, Waseem, and R. Kumar. "A Study on Positive and Negative Effects of Social Media on Society." 30 Oct. 2017, https://www.researchgate.net/profile/Waseem-Akram-19/publication/323903323_A_Study_on_Positive_and_Negative_Effects_of_Social_Media_on_Society/links/5ab1c064a6fdcc1bc0bfefef/A-Study-on-Positive-and-Negative-Effects-of-Social Media-on-Society.pdf?forcedefault=true&__cf_chl_tk=lehWyUcZicXA1l_3TnBYGsEcHN.TYtsx7Vimr2l6hxc-1777123944-1.0.1.1-vsugvMb0cgT2I.v5MMQpLZ8EaMOFQfelRb.69lqmzJE.

Beabadoobee. I-D Meets: Beabadoobee. i-D, 2022, https://youtu.be/B82aGB7KGsU?si=KNK7OIUHEle-s8vD https://www.youtube.com/watch?v=RB6ziiTvbw4&feature=youtu.be.

Belcik , Nathan. "Parental Advisory Explicit Lyrics: A Case Study of Music Censorship and Suppression in America ." Portland State University , 26 Apr. 2012, https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1039&context=younghistorians.

Chetan, Adrian. "Visual Promotion in the Music Industry: Examining the Influence of Visual Branding and Brand Communication on Music Perception ." ResearchGate, 2025, https://www.researchgate.net/profile/Adrian-Chetan/publication/397839990_Visual_promotion_in_the_music_industry_Examining_the_influence_of_visual_branding_and_brand_communication_on_music_perception/links/692087adf4878b75fc783e49/Visual-promotion-in-the-music-industry-Examining-the-influence-of-visual-branding-and-brand-communication-on-music-perception.pdf.

Coulter, Andi. "Marketing Agile Artists: How Music Labels Can Leverage TikTok's Virality." ResearchGate, 2022, https://www.researchgate.net/publication/365112328_Marketing_Agile_Artists_How_Music_Labels_Can_Leverage_TikTok's_Virality.

Crăcană , Sultănica, et al. "The Impact of Social Media on Landscape Businesses." http://www.lsma.ro/index.php/lsma/article/view/2725.

Eminem. "Eminem. Zane Lowe. Part 2." BBC, 19 Nov. 2013, https://www.youtube.com/watch?app=desktop&v=tK9jX91XuT8&pp=ugUEEgJlbg%3D%3D.

Formica , Emily. "Musical Resurgence Phenomenon: A Look Into How Older Songs Are Making Comebacks After Being Featured in Acclaimed Television Shows and Films." University of New Haven , 21 Apr. 2025, https://digitalcommons.newhaven.edu/cgi/viewcontent.cgi?article=1094&context=honorstheses.

Gottfried, Jeffery, and Eugenie Park . Americans' Social Media Use 2025. Pew Research Center, https://www.pewresearch.org/internet/2025/11/20/americans-social-media-use-20 25/.

Hemsley, Jeff. "Virality: Developing a Rigorous and Useful Definition of an Information

Diffusion Process." 2011, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3129424.
Ita, Dara-Abasi. Instagram Explained: Evolution, Features, and Success Story. 27 Feb. 2026, https://www.investopedia.com/articles/investing/102615/story-instagram-rise-1-photo0sharing-app.asp.
Jones, Abby. Samia Rushes Out Stripped-Down "Pool" After Viral Tiny Desk Performance. STEROGUM, 5 Aug. 2025, https://www.google.com/url?q=https://stereogum.com/2318199/samia-rushes-out stripped-down-pool-after-viral-tiny-desk-performance/music&sa=D&source=editors&ust=1777570924225352&usg=AOvVaw0NiK7ufenvO86ICchDpHLB.
Kalia, Ammar. Tiktok- Jazz Sensation Laufey: "It's No Longer About Genre, It's About Feeling and Mood." The Guardian, 3 Mar. 2024, https://www.theguardian.com/music/2024/mar/03/laufey-jonsdottir-iceland-bewitched-tour-grammys-interview.
Kaya, Maria, et al. "How Can Entrepreneurial Musicians Use Electronic Social Networks to Diffuse Their Music." ResearchGate, Jan. 2010, https://www.researchgate.net/publication/277053973_How_can_entrepreneurial_musicians_use_electronic_social_networks_to_diffuse_their_music#read. Keery, Joe. Djo Talks Creating "The Crux," Success of 'End of Beginning' & "Stranger Things" Season 5. Zach Sang Show, 3035, https://m.youtube.com/watch?v=mmUyJZG-jPg https://m.youtube.com/watch?v=ZOUR2jWAkQI&t=2s.
Kentonen, Tiia. "Independent Artists in the Music Industry: What Drives Them and How They Market Themselves." 2018, https://www.theseus.fi/bitstream/handle/10024/150096/thesis-final%20version.pdf?sequence=1&isAllowed=y.
Majka, Marcin. "Understanding Qualitative Analysis in Trend Analysis." ResearchGate, Apr. 2024, https://www.researchgate.net/profile/Marcin-Majka-2/publication/381292669_Understanding_Qualitative_Analysis_in_Trend_Analysis/links/6665f69fde777205a31f6de5/Understanding-Qualitative-Analysis-in-Trend-Analysis.pdf.
Mitski. Mitski Performs a Heavenly Set From '. Raina Douris, 2026, https://www.youtube.com/watch?v=oPL0W0-6seY&feature=youtu.be.
Prakash Yadav , Gyan, and Jyotsna Rai. The Generation Z and Their Social Media Usage: A Review and a Research Outline. ResearchGate, June 2017, https://www.researchgate.net/publication/318005826_The_Generation_Z_and_their_Social_Media_Usage_A_Review_and_a_Research_Outline.
SI, Sajid. "Social Media and Its Role in Marketing." 2016, https://repository.embuni.ac.ke/server/api/core/bitstreams/8295f0fa-d0ad-42be-8082-65dbafeb8e1b/content.

"Social Media Statistics Details." Undiscovered Maine, UMaine, https://umaine.edu/undiscoveredmaine/small-business/resources/marketing-for-small-business/social-media-tools/social-media-statistics-details/ .Accessed 29 Apr. 2026.

The 73% Report. Record Union, https://darenbanarse.com/wp-content/uploads/2025/11/Record_Union-The_73_Percent_Report.pdf.

"The Evolution of Social Media: How Did It Begin, and Where Could It Go Next?" Maryville University Online, 28 May 2020, https://online.maryville.edu/blog/evolution-social-media/.

Thurman, Neil, et al. "My Friends, Editors, Algorithms, and I: Examining Audience Attitudes to News Selection." Digital Journalism, vol. 7, no. 4, Apr. 2019, pp. 447–69. DOI.org (Crossref), https://doi.org/10.1080/21670811.2018.1493936.

Rozehnal, Aleš. "The Role of Social Media in Shaping Society." 2021, https://real.mtak.hu/134573/1/The%20Impact%20of%20Digital%20Platforms%20and%20Social%20Media%20-%20Chapter%207%20-%20Rozehnal.pdf. Wert, Brittani. "Fandom in the Digital Age: Examining Parasocial Relationships Between Fans and Music Artists on Instagram ." Portland State University , 2021, https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=2371&context=honorstheses.

Winder, Parkes. "Tension in the Industry: An Analysis of the Conflict Between Recording Artists and Record Labels Over the Rights To Music Ownership." 13 May 2020, https://repositories.lib.utexas.edu/server/api/core/bitstreams/332f4e44-3643-4e16-9cde-9622a63bee2d/content.

A Comparative Study of Youth Political Agency and the Socio-Cultural Barriers to Legislative Power in Norway and South Korea By Myeounguk Kim

Abstract

Governments worldwide have created youth councils to formalize young people's political voice, but these institutions vary sharply in whether they merely advise or actually decide. This paper asks how cultural perceptions of youth maturity and seniority-based social structures in East Asia shape the institutional design and legislative limits of youth participation bodies, compared with youth council models elsewhere. Using two independent variables — a body's statutory authority (binding versus advisory) and the strength of surrounding age-hierarchy norms — the paper compares Norway, Colombia, Uganda, and South Korea to explain variation in youth legislative agency, defined as the capacity of youth proposals to become law without adult gatekeeping. It finds that Norway's youth councils remain formally advisory but function effectively within a low-power-distance civic culture, while Colombia and Uganda have granted genuinely binding authority through deliberate statutory design, showing that such power is not uniquely "Western." South Korea, by contrast, has built an elaborate consultative architecture that never crosses into binding authority, a pattern this paper attributes to Confucian-derived age hierarchy and an academically saturated model of adolescence. The paper concludes that youth legislative agency depends on the interaction of institutional design and cultural context, and that closing South Korea's agency gap requires a deliberate institutional choice, not merely additional consultation.

Introduction

Over the past three decades, governments across the world have established youth councils, youth parliaments, and youth assemblies with the stated aim of providing young people with a formal role in policymaking processes. Norway implemented this model at the municipal and county levels, while Colombia and Uganda incorporated youth participation into national statutes. In South Korea, a comprehensive network of youth advisory bodies has been developed, extending from the presidential level to individual schools. However, these institutions differ significantly in their functions and authority. In some cases, youth bodies can pass resolutions that public agencies are legally obligated to implement, whereas in others, their role is limited to submitting opinions that may be disregarded by adult authorities. This paper contends that although the formalization of youth participation represents an important step toward democratic inclusion, the extent to which youth bodies are endowed with legislative authority—defined as the capacity to make binding decisions rather than merely issue advisory recommendations—is shaped primarily by the prevailing sociocultural context. In the case of South Korea, a socio-cultural environment characterized by strong age-based hierarchy and an emphasis on academic credentialing over youth autonomy has resulted in youth participation institutions that are procedurally advanced but limited to an advisory function, in contrast to the binding authority achieved by comparable institutions in Colombia and Uganda.

This paper is guided by the following research question: In what ways do cultural perceptions of youth maturity and seniority-based social structures in East Asia shape the institutional design and legislative constraints of youth participation bodies, in comparison to youth council models in other regions? To address this question, the analysis focuses on two independent variables. The first is institutional, referring to whether a youth body's statutory authority is binding or merely advisory. The second is socio-cultural, specifically the degree to which age-hierarchy norms and seniority-based expectations are embedded within a given society. The dependent variable is defined as youth legislative agency, understood as the extent to which youth-generated proposals can be translated into law or policy without being subject to adult oversight or veto. The structure of the paper is as follows. First, it presents a conceptual framework for understanding how institutions with similar formal characteristics may operate differently depending on their cultural context. Second, it examines Norway as a reference case for the Northern European model. Third, it analyzes the cases of Colombia and Uganda, which demonstrate that binding youth authority is attainable outside the Nordic context and thus challenge a simplistic West-versus-rest dichotomy. Fourth, it provides an in-depth analysis of South Korea, arguing that the country's extensive youth participation infrastructure is constrained by cultural norms rooted in Confucian age hierarchy and an academically stratified transition to adulthood. The paper concludes that both formal institutional arrangements and informal cultural norms interact to shape youth agency, and that addressing the limitations of youth legislative authority in South Korea will require changes in both domains.

Theoretical Framework

Institutional theory suggests that what matters for youth political power is the law: if a country passes a statute giving youth councils a formal vote, youth will have power, and if it does not, they will not. This paper takes a more nuanced view, treating formal institutions and informal cultural norms as mutually reinforcing rather than separable. Formal rules matter enormously — a youth council with a legal right to co-decide budget allocations is structurally different from one that can only submit a report that a ministry may set aside. But formal rules are written and interpreted by people who carry cultural assumptions about who deserves to be heard and whose input can be safely disregarded. When a society's dominant cultural script treats youth as unfinished adults — as people still "in training" for citizenship rather than as citizens in their own right — even a well-designed institution will tend to be implemented in a watered-down, consultative way. Conversely, a society that treats a sixteen-year-old's political judgment as roughly commensurate with an adult's is more likely to both pass binding statutes and enforce them.

This dynamic is closely related to what sociologists of childhood and youth studies call "age-grading": the process by which societies divide the life course into stages and attach different rights, responsibilities, and social expectations to each stage. In some societies, adolescence is treated primarily as a preparatory phase oriented toward examinations, credentials, and eventual entry into hierarchical institutions such as universities, firms, or the

military — a phase in which political voice is considered premature. In others, adolescence is treated as an early but legitimate phase of civic membership, in which contributing an opinion to local governance is not fundamentally different in kind from an adult doing so, even if the opinion carries less experience behind it. South Korea, this paper argues, exemplifies the first pattern; the Nordic countries — despite their own limitations, discussed below — lean closer to the second.

Another useful concept is Geert Hofstede's notion of "power distance," which measures the degree to which a culture accepts and expects unequal distribution of power between people of different status (Hofstede). Countries high in power distance tend to treat deference to elders, superiors, and credentialed authorities as a basic social virtue, while countries low in power distance treat challenging or contributing alongside a superior as unremarkable. South Korea has been widely characterized in cross-cultural and organizational research as retaining strong hierarchical expectations organized around age and rank even as the country has modernized economically (Yi et al.). Scandinavian countries, by contrast, sit toward the lower end. This paper does not treat power distance as a deterministic law of culture — Korean society is changing, and Norwegian institutions are not perfectly egalitarian in practice — but as a useful lens for understanding why the same legal template, a youth council with the right to submit proposals, produces different lived outcomes in different places.

Case One - Norway and the Nordic "Consultative-Mandatory" Model

Norway is frequently cited as the leading example of youth political integration, and for good reason: since a 2019 revision of the Local Government Act, Norwegian municipalities and counties have been legally required to establish formal youth councils (National Policies Platform, "5.3"). These councils are composed of members who must not have turned nineteen at the time of their election; they elect their own chairs, and — critically — they have the right to formally address the municipal or county council before it makes decisions on matters that affect young people (National Policies Platform, "5.3"). By 2017, eighteen of Norway's nineteen counties had already established such mechanisms even before they became legally mandatory, reflecting a broader civic culture in which strong, well-resourced youth organizations have operated for decades as recognized partners in policy dialogue rather than as symbolic add-ons (TARSC).

It is important, however, to resist an overly romantic reading of the Norwegian model, because doing so actually sharpens rather than weakens the comparative argument. Norway's youth councils are legally mandatory in the sense that municipalities must create them, but their substantive function remains consultative rather than binding. The councils have "the right to express an opinion" before local government bodies make decisions, not the right to co-decide those outcomes (National Policies Platform, "5.4"). At the national level, the picture is even more limited. Norway experimented with a national "Youth National Assembly" in 2010, and the young delegates who participated voted in favor of creating a permanent state-level youth political body — but a government expert committee reviewing the pilot in 2011 recommended

against institutionalizing it, arguing that regional bodies needed to mature first, and no national binding youth chamber exists today (National Policies Platform, "5.3"). Involvement of youth organizations in national policymaking remains, in the government's own description, consultative and not legally binding (National Policies Platform, "5.4").

This nuance matters for the paper's argument in two ways. First, it shows that even in a country widely regarded as the gold standard for youth civic inclusion, the leap from mandatory existence to binding legislative authority has not been fully made — which suggests that "binding power" is a rarer and more demanding institutional achievement than the mere proliferation of youth councils might suggest. Second, it shows that what truly distinguishes Norway is not that its youth councils vote on binding legislation, but that its socio-cultural environment treats the consultative opinions those councils produce as carrying real practical weight. Norwegian municipal councils that ignore a youth council's formally submitted opinion face real reputational and political costs in a low-power-distance civic culture that expects young people's voices to be taken seriously as a matter of course (TARSC). The institution is advisory, but the surrounding culture makes advisory input function almost as if it were binding in practice. This is a critical insight: culture can substitute for, or substantially reinforce, formal legal authority.

Case Two: Outside the Nordic Context — Colombia and Uganda

If the argument were simply that "Western" democracies grant youth binding power while non-Western ones do not, it would collapse quickly against the evidence; two of the clearest examples of statutorily binding youth authority come from Latin America and Sub-Saharan Africa.

Colombia's Statute of Youth Citizenship (Ley 1622 of 2013, substantially reformed by Ley 1885 of 2018) created a nationwide, multi-tiered system of Municipal, Departmental, and National Youth Councils, whose members — called consejeros — are directly elected by young citizens between the ages of fourteen and twenty-eight (Departamento Nacional de Planeación; Radio Nacional de Colombia). The first-ever national elections for these councils were held in 2021, and subsequent elections have continued on a four-year cycle, most recently in October 2025 (Radio Nacional de Colombia). Colombian youth councilors are not merely permitted to submit suggestions; the reformed statute explicitly tasks municipal council members together with national and local government authorities to work with youth councils on the allocation of public funds, and — in a provision directly relevant to this paper's dependent variable — agreements reached between the government and the youth councils through designated coordination commissions are legally binding (Goitom). This is a meaningfully different institutional design from Norway's "right to express an opinion": Colombian youth councils have a statutory pathway by which their negotiated positions become enforceable commitments rather than optional input.

Uganda offers an even starker example. Under the National Youth Council Act, decisions made by village-level youth councils are explicitly binding on all youths residing in that village,

and the law goes so far as to bind non-citizen youth residents to those decisions as well, even though they cannot vote in them ("National Youth Council Act"). At the national level, Uganda reserves five seats in its Parliament specifically for youth representatives, who are selected through a National Youth Delegates Conference rather than through the ordinary electoral process used for other Members of Parliament ("National Youth Council Act"; Inter-Parliamentary Union). This is a form of guaranteed legislative agency that goes beyond either the Norwegian or the South Korean model: rather than merely consulting a youth body and hoping legislators listen, the Ugandan system reserves actual seats within the lawmaking body itself.

These two cases complicate a simplistic geographic narrative. They show that binding youth authority is not a uniquely "Western" cultural achievement, but rather a product of specific political choices — often made in post-conflict or democratizing contexts where mobilizing youth as a legitimate political constituency served an incumbent government's broader legitimacy-building project. Colombia's youth statute emerged in a country working to include a generation shaped by decades of armed conflict; Uganda's youth quota emerged from a political settlement in which the ruling movement sought to formally represent a demographically dominant youth population. In neither case did binding youth power require the absence of hierarchical culture — Uganda, in fact, retains strong elder-based traditional authority structures in many regions. This suggests an important qualification to the paper's socio-cultural argument: cultural hierarchy is a powerful constraint, but it is not insurmountable, and institutional design choices can partially override hierachy it when political incentives align.

Case Three: South Korea's Advisory Architecture and Its Cultural Foundations

South Korea presents what might be called a paradox of participatory sophistication without legislative agency. The country has built an unusually dense and well-funded architecture of youth participation. The National Youth Council (청소년특별회의), established in 2004 under a presidential pledge and operating under the Ministry of Gender Equality and Family, brings together youth members selected centrally alongside hundreds more organized at the metropolitan and provincial level (Ministry of Gender Equality and Family). A further layer of Youth Participation Committees and Youth Steering Committees operates in local governments and individual youth facilities, coordinated through a national online portal that allows young people to submit and vote on policy proposals (Korea.net; Open Government Partnership). In 2025, the government further announced plans to expand a national youth policy committee with new expert subcommittees and to appoint more young representatives to government councils as part of a second five-year master plan for youth policy running through 2030 (Jung). On paper, this is one of the most elaborate state-sponsored youth participation infrastructures in the world.

Yet the defining institutional feature of the Korean system is that it channels youth input toward the executive branch for discretionary consideration rather than toward the legislature for binding enactment. South Korea's Youth Working Group, for instance, was created specifically

to “convey youth perspectives” to government ministries — a framing that describes influence over deliberation, not a share of decision-making authority (Open Government Partnership). There is no mechanism, comparable to Colombia’s binding coordination-commission agreements or Uganda’s reserved parliamentary seats, by which a youth proposal automatically becomes policy or law; it is a persuasive input into an adult-controlled decision, not a co-equal decision in itself. The 2025 policy direction reinforces rather than resolves this pattern: it promises to “expand” committees and “appoint more” youth representatives to existing councils, language that describes an intensification of consultation, not a transfer of decision-making authority (Jung). Consistent with this pattern, one recent account of Korean youth civic life describes young South Koreans as increasingly distrustful of elected politicians and turning to independent civil-society projects precisely because they do not feel their input through official channels translates into real decision-making power (Vidwans). Why has South Korea built such an elaborate consultative machine without ever crossing the threshold into binding authority, when Colombia and Uganda — countries with considerably less developed state administrative capacity — made that leap over a decade ago? This paper’s central claim is that the answer lies substantially in South Korea’s socio-cultural landscape, and specifically in the enduring strength of Confucian-derived age hierarchy and the country’s intensely credential-oriented pathway to adult social status.

The Confucian influence on Korean social organization is well documented. Age itself has long conferred social authority in Korean life largely independent of an individual’s personal merit, and observers of contemporary Korean society continue to describe hierarchy, respect for age, and seniority as active, reproduced features of family, workplace, and social life rather than fading traditional relics (“South Korea: Daily Life”). A 2008 newspaper analysis of Korean workplace and school culture similarly described seniority-based structures as both a historical catalyst for national development and a persistent source of friction between merit-based and age-based logics of authority (“Seniority-Based Culture”). A peer-reviewed study in Pacific Affairs similarly found that seniority and hierarchy continue to shape Korean labor-market behavior despite decades of economic modernization, distorting outcomes for both older and younger workers (Yun). Even within Korean higher education — an environment ostensibly organized around academic merit — a 2023 peer-reviewed study found that a Confucian-influenced age and rank hierarchy restricts opportunities for lower-status junior members, including undergraduates and junior graduate students, to take on visible, authorship-bearing research roles, reserving those roles for senior, credentialed figures (Yi et al. 450). If this pattern holds even within universities, where meritocratic norms are supposed to be strongest, it should be unsurprising that a similarly graded logic extends into the design of youth political institutions.

The cultural weight of age in Korean society is not just anecdotal; it has recently been the subject of national legislation. In December 2022, South Korea’s National Assembly passed laws abolishing the traditional “Korean age” counting system — under which a person is considered one year old at birth and gains a year every January 1 regardless of actual birth date — in favor

of the international standard, with the change taking legal effect in June 2023 (Park). Coverage of the change quoted a Korean anthropologist explaining that "age really matters" in South Korean culture because it affects a person's relative social status and dictates which honorific forms of address one must use toward others — a dynamic that persisted even as the government moved to simplify the legal calculation of age itself ("Why Are South Koreans"). The legislation is itself an evidence of how deeply age-consciousness is embedded in ordinary Korean social interaction, reinforcing the claim that youth political institutions in Korea operate inside a broader cultural system in which one's age routinely determines whose voice carries authority.

This cultural pattern intersects with a closely related feature of Korean youth life: the near-total organization of adolescence around a competitive academic trajectory oriented toward university entrance and, eventually, employment at a reputable profession. Where Norwegian civic culture treats a sixteen-year-old showing up to a municipal council meeting as an ordinary act of local citizenship, the dominant script of Korean adolescence treats the years before university entrance as a period properly devoted to studying rather than governing — a script that both youth and adults have strong incentives to internalize, since academic performance during this period disproportionately determines lifetime social status. Under such conditions, even a young person who is personally motivated to engage in youth councils faces a real opportunity cost, and society at large has been culturally primed to view youth political engagement as a supplement to, rather than a partial substitute for, academic preparation.

The consequence, this paper argues, is that South Korea's National Youth Council and its subordinate bodies function less as instruments of youth legislative agency and more as sophisticated simulations of participation: structured, well-resourced, digitally enabled, and genuinely capable of surfacing youth policy preferences to decision-makers — but ultimately dependent on adult executive discretion. This is not a matter of legal drafting failure. South Korea has demonstrated ample administrative capacity to design far more binding institutions if it chose to, as its sophisticated bureaucracy in other domains attests (Jung). Rather, it reflects a socio-cultural equilibrium in which formally empowering youth to make binding decisions over adults — even decisions confined to youth policy specifically — would cut against a resilient hierarchical common sense about whose judgment is authoritative.

Comparative Analysis

A comparative analysis of the four cases clarifies the relationship between the two independent variables and the dependent variable of youth legislative agency. Uganda and Colombia both demonstrate high scores on the institutional variable, as each country grants genuinely binding authority through mechanisms such as reserved legislative seats or enforceable coordination agreements. Correspondingly, both cases exhibit comparatively high levels of youth legislative agency in practice, despite facing significant governance challenges. In contrast, Norway scores lower on the institutional variable, since its councils remain consultative even where their establishment is mandatory. However, Norway still demonstrates a meaningfully high degree of functional youth agency. This outcome can be attributed to its

low-power-distance civic culture, which ensures that consultative input from youth is given substantial practical weight and imposes reputational costs if disregarded. South Korea presents a different pattern, combining a moderate-to-high institutional score, characterized by an elaborate, well-funded, multi-tiered consultative architecture, with a high-power-distance, hierarchically organized socio-cultural environment. As a result, South Korea records the lowest level of youth legislative agency among the four cases, as neither variable independently compensates for the weakness of the other. These findings indicate that youth legislative agency is not determined solely by legal design or by cultural openness, but rather by the interaction between institutional mechanisms and socio-cultural context. Strong institutions can partially compensate for a less youth-friendly culture, as observed in Uganda, while a robust civic culture can partially offset weaker formal authority, as seen in Norway. However, in cases where both a binding institutional mechanism and a low-hierarchy cultural environment are absent, as in South Korea, the outcome is likely to be participatory but not legislative in nature.

Discussion and Limitations

Several important limitations should be noted. First, this comparison is based on a limited number of cases and should be considered illustrative rather than a definitive test of a general theory. A more comprehensive research design would incorporate additional East Asian cases, such as Japan and Taiwan, as well as further Nordic or Latin American cases, to assess whether the observed patterns are consistent across each regional cluster. Second, culture is dynamic, and the characterization of Korean age-hierarchy norms in this paper reflects a still-dominant yet increasingly contested pattern. Korean society has demonstrated a willingness to reconsider even deeply rooted age-based practices, as evidenced by the 2022–2023 legislative overhaul of the traditional age-counting system, although the cultural significance attached to age has proven more persistent than the counting convention itself (Park). Third, the cases of Uganda and Colombia indicate that binding youth authority can emerge even in the presence of hierarchical cultural norms. Therefore, South Korea's progress toward greater youth legislative agency does not require a complete cultural transformation. A sufficiently deliberate institutional intervention, similar to those implemented in Colombia and Uganda, could plausibly establish binding youth authority within a hierarchical cultural context, provided there is adequate political will.

Conclusion

This paper has argued that formal institutional design and informal socio-cultural norms jointly determine whether youth participation bodies translate into genuine legislative agency. Norway demonstrates that even primarily consultative institutions can function effectively when embedded in a low-power-distance civic culture that treats youth input as carrying real weight. Colombia and Uganda demonstrate that binding youth authority is achievable through deliberate statutory design even outside the Nordic or “Western” context, complicating any simple civilizational narrative about who is capable of empowering youth. South Korea, by contrast, has built one of the world’s most elaborate consultative youth participation architectures without ever

crossing the threshold into binding legislative authority, and this paper has argued that this outcome is substantially explained by the continued strength of Confucian-derived age hierarchy and an academically saturated model of adolescence that together position youth political voice as something to be solicited and considered by adults, rather than something youth hold as an independent legislative entitlement. Addressing this gap would require not merely additional consultative committees, of which South Korea already has an abundance, but a deliberate institutional choice, of the kind Colombia and Uganda have already made, to convert youth policy input into a genuinely binding component of the legislative process.

Works Cited

Choi, Jeong-yoon. “Korea Likely to Keep Criminal Age Threshold at 14.” The Korea Herald, 3 May 2026, www.koreaherald.com/article/10730321.

Cultural Atlas. “South Korean Culture: Family.” SBS, culturalatlas.sbs.com.au/south-korean-culture/south-korean-culture-family. Accessed July 2026.

Departamento Nacional de Planeación. “Sistema Nacional de Juventud.” Gobierno de Colombia, www.dnp.gov.co/LaEntidad_/subdireccion-general-prospectiva-desarrollo-nacional/direccion-desarrollo-social/Paginas/sistema-nacional-de-politicas-publicas-de-juventud.aspx. Accessed July 2026.

Goitom, Hanibal. “Colombia: President Promulgates Reform of Youth Citizenship Statute.” Library of Congress, Global Legal Monitor, 13 June 2018, www.loc.gov/item/global-legal-monitor/2018-06-13/colombia-president-promulgates-reform-of-youth-citizenship-statute/.

Hofstede, Geert. Culture’s Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations. 2nd ed., Sage Publications, 2001.

Inter-Parliamentary Union. “Uganda: Data on Youth.” IPU Parline, data.ipu.org/parliament/UG/UG-LC01/data-on-youth/. Accessed July 2026.

Jung, Min-kyung. “S. Korea’s Policy Blueprint for Youth Focuses on Helping Them Find Their Feet.” The Korea Herald, 22 Sept. 2025, www.koreaherald.com/article/10580713.

Korea.net. “The Ministry of Gender Equality and Family Discusses Youth Policy Direction for an Inclusive Nation.” Press release, www.korea.net/Government/Briefing-Room/Press-Releases/view?articleId=4523&type=O. Accessed July 2026.

Ministry of Gender Equality and Family, Republic of Korea. Official website, www.mogef.go.kr/eng/am/eng_am_f007.do. Accessed July 2026.

National Policies Platform. “5.3 Youth Representation Bodies.” Youth Wiki, European Education and Culture Executive Agency, national-policies.eacea.ec.europa.eu/youthwiki/chapters/norway/53-youth-representation-bodies. Accessed July 2026.

---. “5.4 Young People’s Participation in Policy-Making.” Youth Wiki, European Education and Culture Executive Agency, national-policies.eacea.ec.europa.eu/youthwiki/chapters/norway/54-young-peoples-participation-in-policy-making. Accessed July 2026.

“National Youth Council Act.” Uganda Legal Information Institute, Cap. 319, ulii.org/en/akn/ug/act/statute/1993/2/eng@2023-12-31.

Open Government Partnership. “Engaging Youth in Policy Making (KR0049).” Republic of Korea, www.opengovpartnership.org/members/republic-of-korea/commitments/KR0049/. Accessed July 2026.

Park, Seongryeol. “Koreans Becoming Younger – Unification of Age-Counting Systems.” In Custodia Legis, Law Library of Congress, 27 Jan. 2023, blogs.loc.gov/law/2023/01/koreans-becoming-younger-unification-of-age-counting-systems/.

Radio Nacional de Colombia. “Juventud con voz y voto: así funcionan los Consejos de Juventud en Colombia.” www.radionacional.co/actualidad/politica/que-son-los-consejos-de-juventud-en-colombia-y-por-que-son-importantes. Accessed July 2026.

“Seniority-Based Culture.” The Korea Times, 8 Oct. 2008, www.koreatimes.co.kr/southkorea/society/20081008/seniority-based-culture.

“South Korea: Daily Life and Social Customs.” Encyclopædia Britannica, www.britannica.com/place/South-Korea/Daily-life-and-social-customs. Accessed July 2026.

Training and Research Support Centre. Norway Case Study: Supporting Children’s Voice and Agency in Policy Change on Child Wellbeing. TARSC, 2019, www.tarsc.org/publications/documents/FCH%20Norway%20case%20study%202019.pdf.

Vidwans, Prachi. “South Korea’s Young People Are Fed Up with Party Politics.” Institute of Current World Affairs, 23 May 2024, www.icwa.org/south-korea-young-people-politics/.

“Why Are South Koreans Suddenly Younger? New Age System Explained.” Al Jazeera, 28 June 2023, www.aljazeera.com/news/2023/6/28/why-are-south-koreans-one-year-younger-today.

Yi, Joseph, et al. “Student Research Participation in South Korea.” PS: Political Science & Politics, vol. 56, no. 4, 2023, pp. 449-54.

Yun, Ji-Whan. “The Myth of Confucian Capitalism in South Korea: Overworked Elderly and Underworked Youth.” Pacific Affairs, vol. 83, no. 2, 2010, pp. 237-59.

Optimizing Growth Media for *Acinetobacter Baumannii* using Genome-Scale Metabolic Modeling By Jaden de Venecia

Abstract

Acinetobacter baumannii is an opportunistic bacterial pathogen well known for its ability to adapt to diverse environments and to resist antibiotics. These adaptive abilities are linked to the bacterium's metabolism, which is heavily dependent on both nutrient availability and oxygen use. Consequently, studying how different carbon sources affect growth can help explain how *A. baumannii* adapts metabolically.

In this investigation, I used Flux Balance Analysis (FBA) simulations to predict the effects of changes in gene expression, nutrient availability, and oxygen levels on *A. baumannii*'s growth. The media conditions tested were aerobic (acetate), anaerobic, blood-like glucose, glycerol, and fructose. Furthermore, I conducted gene-deletion simulations and a production envelope analysis to enable the Genome-Scale Metabolic Model (GEM) to identify essential genes and better understand the effect of changes in oxygen uptake on predicted growth.

This paper aims to investigate how different media conditions affect predicted growth, oxygen consumption, and essential genes of *A. baumannii*. The results showed that the glycerol condition had the highest predicted growth rate, the anaerobic condition had no growth rate, and the aerobic condition consumed the most oxygen and had an additional 10 essential genes (used in gluconeogenesis, ATP synthase, and acetate transport). The findings are significant as they deepen our understanding of how different carbon sources and oxygen consumption affected the growth and metabolic behavior of *A. baumannii*. Therefore, through these discoveries, I found that identifying and targeting these essential genes, as well as better understanding how the bacterium grows under different environmental conditions, is crucial for future antimicrobial research.

Introduction

Acinetobacter baumannii is a gram-negative, opportunistic bacterial pathogen commonly found in soil, water, and hospital environments (Wilharm et al., 2026). Classified as a "superbug," certain strains of *A. baumannii* have developed high-level resistance to multiple antibiotics, including carbapenems, thereby rendering them Extensively Drug-Resistant. Consequently, the rising threat of such antimicrobial-resistant (AMR) bacteria prompted the World Health Organization to designate carbapenem-resistant *A. baumannii* strains as a Critical Priority pathogen, underscoring the urgent need for effective treatment options. Without intervention, this pathogen poses a serious risk, especially in hospital settings, where the likelihood of encountering untreatable infections increases significantly.

The history of *Acinetobacter* can be traced back to 1911, when Dutch microbiologist M.W. Beijerinck isolated an organism named "Micrococcus calco-aceticus" from soil, later recognized as part of the early taxonomic history of the *Acinetobacter* genus (Karah et al., 2025). However, it was not until 1986 that Bouvet and Grimont designated *A. baumannii* as a new and distinct species within the genus. This taxonomic development is important as it shows how *A.*

baumannii evolved into a distinct species with clinical importance, emphasizing the need for further biological research into its antibiotic resistance.

In terms of its structure, *A. baumannii* exhibits a unique cell wall structure, which results in a characteristic pink or red stain during the Gram stain procedure (Tripathi et al., 2025). This coloration is due to its thin peptidoglycan layer and the presence of an outer membrane, which contributes to its antibiotic resistance and ability to thrive in different environments (Zhou et al., 2023). Additionally, *A. baumannii* can survive on surfaces for extended periods, making it a persistent threat to hospitalized patients, particularly those with weakened immune systems. This can lead to severe infections, including pneumonia, bloodstream infections, urinary tract infections, wound infections, and other serious health complications.

Given the increasing threat posed by AMR microorganisms, understanding *A. baumannii* metabolism is crucial for developing new treatments against this bacterium. While traditional and experimental laboratory methods can map the bacterium's metabolism, they often require significant time and financial resources. Additionally, the complexity of biological systems can often limit such approaches. By employing GEMs, which mathematically represent organisms' metabolic networks, one can effectively analyze how *A. baumannii* adapts and survives under various environmental conditions without relying heavily on experimental methods.

Over time, multiple GEMs of *A. baumannii* have been created to study the bacterium, each with varying numbers of reactions, metabolites, and genes. For instance, model "iAB5075" contains around 2,229 reactions, 1,530 metabolites, and 1,015 genes (Zhao et al., 2020). However, the specific model I will be studying is "iACB23LX" (Leonidou et al., 2024), the most recent and comprehensive model for strain ATCC 17978, containing approximately 2,321 reactions, 1,660 metabolites, and 1,164 genes. This model was evaluated using the SMBL validator, MEMOTE, and FROG to ensure its biological and structural accuracy (Leonidou et al., 2024). Therefore, this model is ideal for simulating *A. baumannii's* metabolic network under different environmental conditions and predicting gene essentiality.

In this study, I conducted a Flux Balance Analysis (FBA) on GEMs to simulate the effects of changes in gene expression, nutrient availability, and oxygen levels on *A. baumannii*'s growth. Furthermore, I conducted gene-deletion simulations to enable the GEM to identify essential genes and pathways as potential new drug targets. Through these analyses and changes in media conditions, I gained better insight into how this particular strain of *A. baumannii* adapts to different environments and identified essential genes specific to each environment.

Methods

Metabolic Model and Software Used For Flux Balance Analysis (FBA)

The GEM that I have chosen for this study is "iACB23LX", which is of the strain ATCC 17978 and includes 2,321 reactions, 1,660 metabolites, and 1,164 genes (Leonidou et al., 2024). The model was obtained from the BioModels database as an SBML file, allowing me to use COBRApy, a package for genome-scale metabolic modeling and Flux Balance Analysis

(Ebrahim et al., 2013). Additionally, COBRApy was used to import the model ("read_sbml_model()"), simulate different media, run a single-gene deletion analysis ("single_gene_deletion()"), perform production envelope analysis, and generate a summary of the model ("model.summary()"). Lastly, Flux Balance Analysis ("model.optimize()") was also used to predict *A. baumannii*'s growth rate by quantifying metabolite fluxes (Orth et al., 2010). These results were then organized using tables and plotted using matplotlib (Hunter, 2007). For instance, the aerobic medium was loaded using "model = read_sbml_model('iACB23LX.xml')". Then, "model.summary()" was used to present a summary of the model, including the objective growth and uptake reactions. Lastly, after both the model was loaded and a summary was created, all the necessary analyses were run for the condition.

Modeling For Growth In Different Conditions

To predict growth under different environments (aerobic, anaerobic, blood-like glucose, glycerol, and fructose media), I used FBA to calculate the optimal growth rate of *A. baumannii* in each condition by optimizing the model's biomass objective function. For instance, the anaerobic condition was simulated by removing oxygen uptake via "model.medium" and setting the oxygen import to "0.0". To do this, I first created a new "anaerobic medium" by copying and altering the existing "aerobic medium" to ensure that the original was not affected ("anaerobic_medium = aerobic_medium.copy()"). I then found the oxygen reaction ID, "EX_o2_e," and set the exchange reaction to "0.0" ("anaerobic_medium["EX_o2_e"] = 0.0"). Lastly, I applied the anaerobic medium to the model using "model.medium = anaerobic_medium", which enabled the subsequent simulations to run under that condition.

I also conducted simulations in a fructose environment by first creating a new medium dictionary using "fructose_medium". Then, I added fructose as the carbon source by replacing the acetate (the default carbon source) in the reaction. This was done by identifying both exchange reaction IDs: acetate ("EX_ac_e") and fructose ("EX_fru_e"). Next, I set the fructose exchange reaction to "8.33" and added nitrogen, phosphate, and sulfur sources, as they are basic nutrients for growth, therefore helping the model to produce biomass. Lastly, I added both ions and trace nutrients, which help support normal cell processes. For example, chloride and hydrogen ions help maintain chemical balance within the cell. From there, I finally applied the fructose medium to the model using "model.medium = fructose_medium", which enabled the subsequent simulations to run under that condition. Each condition, therefore, was chosen with the utmost intention. While the aerobic medium made oxygen available, the anaerobic medium set the oxygen uptake to "0.0", indicating no oxygen uptake. Lastly, the fructose and glycerol media were chosen to exemplify the effect of changing carbon sources, while the blood-like glucose condition represented a modified nutrient condition.

Furthermore, the uptake rates in the fructose and glycerol media were adjusted to match the same total carbon input as acetate. Acetate has an uptake rate of 25 $mmol\ acetate\ gDW^{-1}h^{-1}$ and contains 2 carbon atoms, with 50 carbon units. To ensure that the results were fair, I adjusted the uptake rates of fructose and glycerol to equal 50 carbon units. For

example, fructose has 6 carbon atoms, which means its uptake rate must be 8.33 $mmol\ glycerol\ gDW^{-1}h^{-1}$. Therefore, using similar procedures across all my selected media, I determined *A. baumannii*'s optimal growth rate in each environment. Then, based on the results, I ran multiple simulations and analyses to provide the best possible environment for *A. baumannii* growth.

Single-gene Deletion

Another simulation I ran was a single-gene deletion analysis to knock out a specific gene *in silico*. The growth rate was recorded after each deletion under separate conditions. For instance, genes that showed no predicted growth after deletion were considered essential. By running this process across all media, I was able to determine whether certain genes were essential only under specific environmental conditions. This enabled me to present the data in both a table and a Venn diagram to gain better insight into the genes shared across media.

For instance, under aerobic conditions, I first loaded the medium and measured the initial growth rate. I then ran a single deletion analysis using "result_aerobic_single_deletion = single_gene_deletion(model,processes=4)". This allowed me to knock out each gene individually *in silico*. Then, after each gene deletion, a new growth rate was predicted to see if the gene was essential for growth. As an example, the simulation output included the line "{YP_001084881_1} 0.550313 optimal," indicating that a certain gene maintained the same growth rate, showing it is non-essential.

Based on the results, I created a table showing which gene was deleted, the predicted post-deletion growth, and whether the optimization was successful. Then, the essential genes were filtered using the code "essential_aerobic = result_aerobic_single_deletion[result_aerobic_single_deletion["growth"] < 1e-6]". This means genes were considered essential if their growth rate dropped below "1e-6." This value was chosen because it is very close to zero, representing nearly no growth. Finally, a few essential genes were presented using the command "essential_aerobic.head()". For example, "{YP_001084005_1} 4.350113e-16 optimal," shows that this gene exhibited growth below "1e-6," which means that COBRApy considers the gene essential.

Production Envelope Analysis

I also ran a production envelope analysis in COBRApy using "prod_env = production_envelope(model, ["EX_o2_e"])," to calculate the maximum possible growth rate for each oxygen uptake level. This analysis was used to examine further the effect of changes in oxygen uptake on predicted growth. Finally, using the results from the production envelope, I used matplotlib (Hunter, 2007) to present the data in a graph and analyze trends for each medium.

Processing Essential Gene Comparisons

When comparing the essential genes, I filtered them using the rule that if growth fell below "1e-6," the gene would be considered essential, as *A. baumannii* would no longer grow without that particular gene. Then, I used the code "essential_aerobic=set(essential_aerobic["ids"].astype(str))" to convert the aerobic essential gene table into a set of gene IDs and compared them with the other essential genes from separate media. The essential genes were then compared using set intersections and set differences.

Additionally, to further analyze the essential genes specific to each condition, the 10 genes essential only in the aerobic medium were identified. Then I used the model's gene-reaction associations to determine which reactions are associated with those 10 genes. Finally, from those reactions, I identified the possible enzymes associated with each gene. For instance, when investigating the gene "YP_001083307_1," I used "gene.id" to obtain the ID, "gene.name" to identify the name, and a "for loop" of the reaction ID, names, and equation to find the necessary data for a particular gene. This process was then repeated for all 10 genes.

Results

Summary of Growth Rate and Oxygen Consumption Across Different Media Conditions

Based on the results below, the model predicted different growth rates depending on the medium used. This pattern was present for all media: aerobic, anaerobic, blood, fructose, and glycerol, each chosen for specific purposes. Firstly, the aerobic (default) condition was chosen to represent *A. baumannii* under oxygenated conditions. In contrast, anaerobic conditions were used to show the bacterium's growth when oxygen is removed. Additionally, I simulated two conditions in which I switched the carbon source from acetate to either fructose or glycerol. Lastly, I did a final condition, the blood-like glucose condition, to simulate the bacterium's growth in a blood-like environment.

Table 1. Growth Rate and Oxygen Consumption in Different Media

Medium	Growth Rate (h^{-1})	Oxygen Consumption ($mmol\ O_2\ gDW^{-1}h^{-1}$)
Aerobic (default)	0.55	26.25
Anaerobic	0.00	0.00
Blood-like Glucose	0.69	20.35
Fructose	0.74	18.22
Glycerol	0.86	21.39

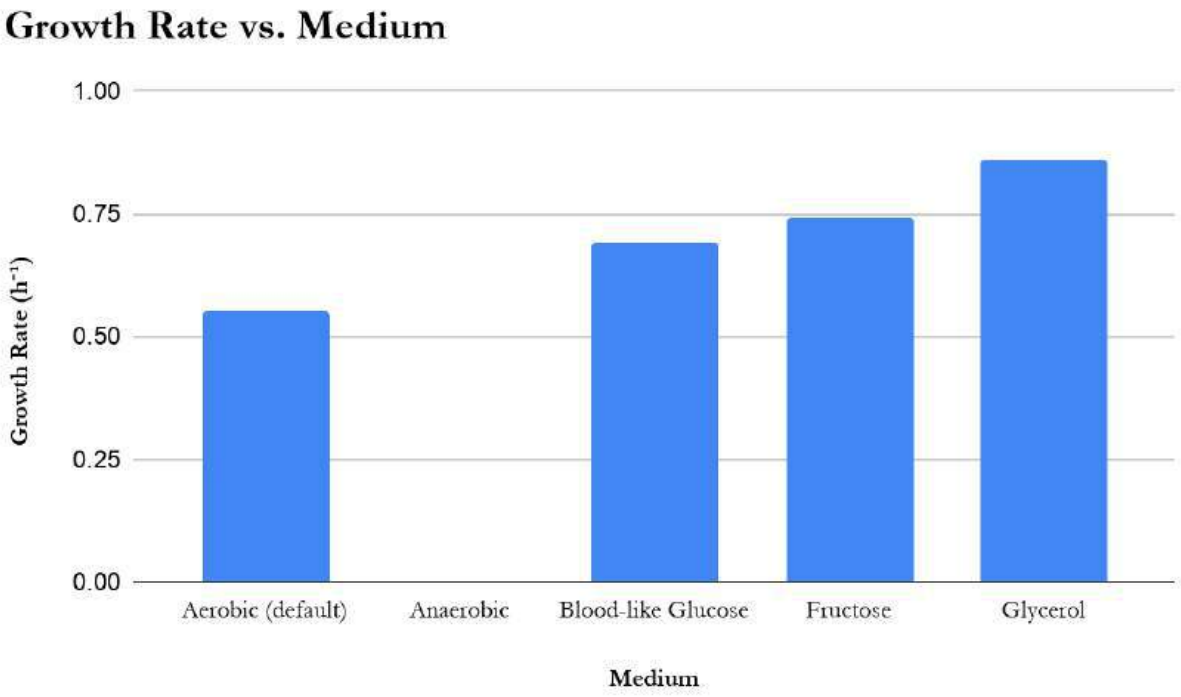


Figure 1. Different Simulated Growth Rates per Media

The results above presented that the aerobic conditions had a growth rate of 0.55 per hour, the anaerobic conditions had no growth, the blood-like glucose medium had a growth rate of 0.69 per hour, the fructose medium had a growth rate of 0.74 per hour, and the glycerol medium had a growth rate of 0.86 per hour (Table 1, Figure 1). Based on these results, it is evident that the glycerol medium produced the highest predicted growth rate, the aerobic medium the lowest, and the fructose and blood-like media relatively higher growth rates across all media. However, the anaerobic conditions resulted in zero growth, indicating that the model could not produce biomass in the absence of oxygen uptake.

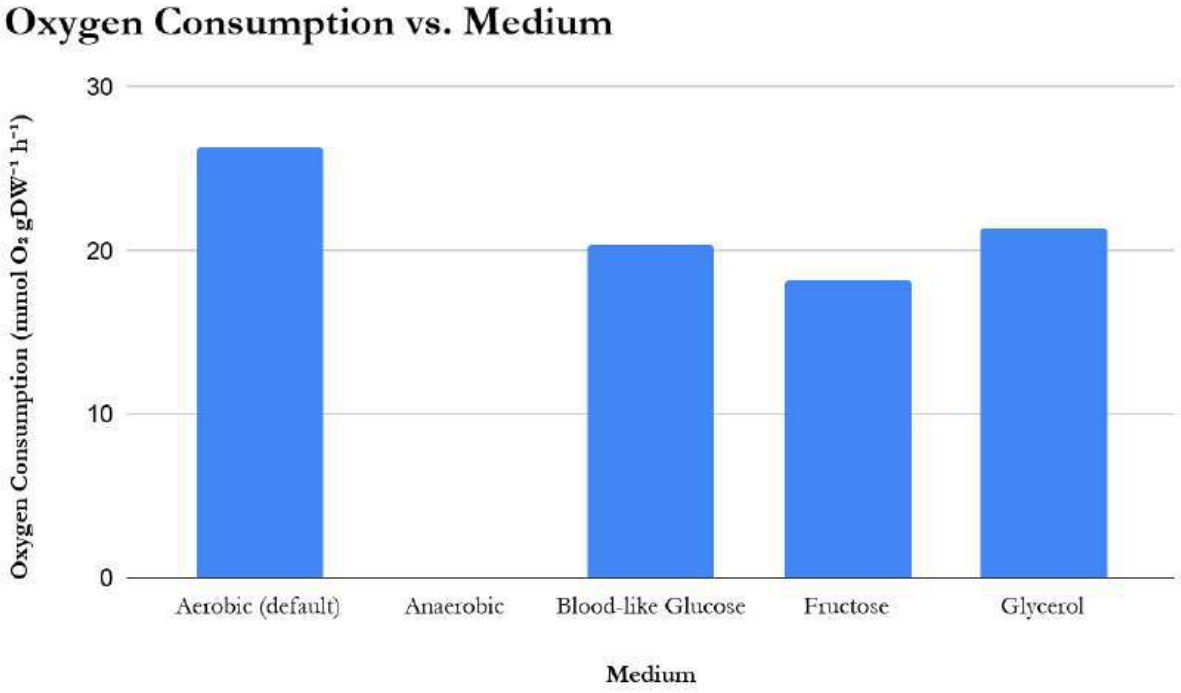


Figure 2. Different Oxygen Consumption Values per Media

When looking at the oxygen consumption results across the media, I found that the aerobic condition had the highest oxygen consumption at 26.25 $mmol\ O_2\ gDW^{-1}h^{-1}$. In contrast, the glycerol and blood-like media had a lower oxygen consumption rate of 21.39 and 20.35 $mmol\ O_2\ gDW^{-1}h^{-1}$, respectively, despite a predicted higher growth rate. Finally, fructose had the lowest oxygen consumption among the different media at 18.22 $mmol\ O_2\ gDW^{-1}h^{-1}$. As for the anaerobic condition, oxygen consumption was 0, indicating no simulated growth (Table 1, Figure 2).

Based on the results, it is evident that higher oxygen consumption does not always mean higher growth. This is exemplified by the glycerol medium, which had the highest growth rate among the media, yet its oxygen consumption was lower than under aerobic conditions. Additionally, glycerol and fructose produced more growth than the aerobic medium, suggesting that high oxygen consumption does not necessarily indicate efficient conversion of the carbon source into biomass. Furthermore, glycerol was the most effective medium, as it had the highest predicted growth rate and lower oxygen consumption than the aerobic condition (which uses acetate), indicating it served as a more effective carbon source for growth. As for the blood-like and fructose media, they both still supported growth, yet not as strongly as the other conditions. Lastly, the absence of growth under anaerobic conditions suggests that oxygen is required for *A. baumannii* to grow. Therefore, the data show that, because the carbon input was equal across the

media, the differences in growth and oxygen consumption resulted from how the model allocated each carbon source to growth. Hence, it indicates that while oxygen availability is necessary for growth, the type of carbon source more strongly influences how efficiently that oxygen is used to produce biomass.

Essential Gene Counts and Overlaps Across Different Media Conditions

The results shown below are the essential gene counts across the different media conditions. These results helped me identify similarities and differences in the genes the model uses to produce biomass across different environments. To determine an essential gene, I used the rule that if growth fell below "1e-6," the gene would be considered essential, as *A. baumannii* would no longer grow without that particular gene. Then, following this rule, I was able to determine which media conditions required more or fewer genes to survive.

Table 2. Essential Gene Counts per Medium

Medium	**Essential Genes**
Aerobic (default)	139
Anaerobic	N/A
Blood	129
Fructose	129
Glycerol	129

When comparing essential gene counts across the selected media, the aerobic condition had 139 predicted essential genes. In contrast, the other media had fewer (129), suggesting that these media required fewer genes to sustain growth than the aerobic condition, which required more reactions and pathways. Lastly, the anaerobic condition showed no essential genes, as there was no growth, indicating that *A. baumannii* could not survive without oxygen (Table 2).

Additionally, upon close examination, I found many similarities and differences among the conditions. Under the aerobic condition, the bacterium had a growth rate of 0.55 per hour, an oxygen consumption of 26.25 $mmol\ O_2\ gDW^{-1}h^{-1}$, and an essential gene count of 139 (Tables 1 and 2). This shows that the bacterium required more genes to survive under this condition. On the other hand, I found the glycerol medium to be the most effective for growth, despite its lower oxygen consumption. This was because it exhibited the highest predicted growth rate of 0.86 per hour, an oxygen consumption of 21.39 $mmol\ O_2\ gDW^{-1}h^{-1}$, and 129 essential genes (Tables 1 and 2). This indicates that, even with lower oxygen consumption, it was the most effective medium due to the high growth rate and fewer essential genes for growth.

Additionally, the fructose medium supported growth but yielded a lower predicted growth rate than the glycerol medium. It had a growth rate of 0.74 per hour, an oxygen consumption of 18.22, and an essential gene count of 129 (Tables 1 and 2). This shows that while the condition supported growth with lower oxygen consumption, it was still less efficient for growth than glycerol, indicating that higher oxygen consumption does not necessarily imply a higher growth rate. Finally, under anaerobic conditions, there was no growth, no essential genes, and no oxygen consumption, as the bacterium was unable to grow (Tables 1 and 2). This shows that *A. baumannii* requires oxygen for growth, as evidenced across all media.

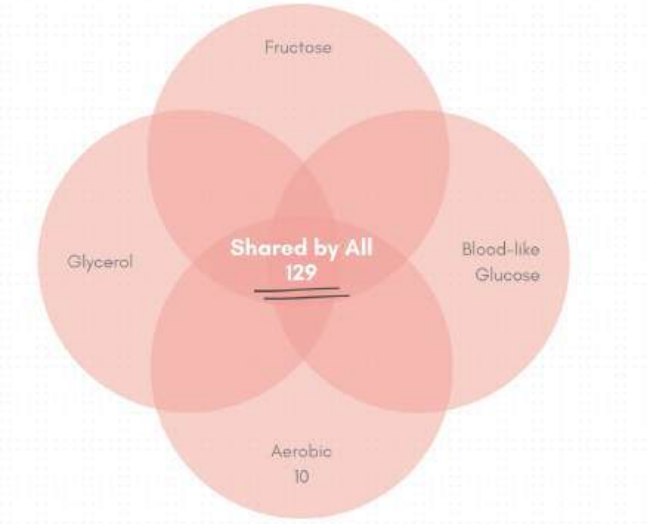


Figure 3. Venn Diagram Representing the Overlap of Essential Gene Counts per Medium

Using the matplotlib library (Hunter, 2007), I was able to count the number of genes shared by all four growing media and by each medium. The results indicate that 129 essential genes were shared across all four growing media, and 10 genes were unique to the aerobic medium (Figure 3). This finding of 10 genes specific to aerobic conditions further supports the hypothesis that *A. baumannii* requires more genes to survive in this medium. To further investigate this, I found the 10 unique reactions of the aerobic medium.

Unique Reactions Performed by Aerobic Medium

To further investigate the 10 genes unique to the aerobic condition, I created a table outlining each gene's ID and reaction(s). Then, from the results, I was able to split up the genes into three different groups: supports glycolysis, supports acetate transport, and supports ATP production. The group of genes that support glycolysis/gluconeogensis includes: "YP_001083307_1", "YP_001084572_1", "YP_001085521_1", and "YP_001084927_1". The gene that supports acetate transport is "YP_001086291_1." Finally, the group of genes that support ATP production include: "YP_001083233_1", "YP_001083241_1", "YP_001083239_2", "YP_001083238_2", and "YP_001083240_2." Below is the table that provides more detail on the genes:

Table 3. Essential Genes of Aerobic Condition

10 Genes Unique to the Aerobic Medium	Essential Gene Data
YP_001083307_1	Reaction ID: PGM Reaction Name: Phosphoglycerate mutase Reaction Equation: 2pg_c <=> 3pg_c
YP_001084572_1	Reaction ID: PGK Reaction Name: Phosphoglycerate kinase Reaction Equation: 3pg_c + atp_c <=> 13dpg_c + adp_c
YP_001083240_2	Reaction ID: ATPS4rpp Reaction Name: ATP synthase (four protons for one ATP) (periplasm) Reaction Equation: adp_c + 4.0 h_p + pi_c <=> atp_c + h2o_c + 3.0 h_c
YP_001085521_1	Reaction ID: GAPD Reaction Name: Glyceraldehyde-3-phosphate dehydrogenase Reaction Equation: g3p_c + nad_c + pi_c <=> 13dpg_c + h_c + nadh_c
YP_001083238_2	Reaction ID: ATPS4rpp Reaction Name: ATP synthase (four protons for one ATP) (periplasm) Reaction Equation: adp_c + 4.0 h_p + pi_c <=> atp_c + h2o_c + 3.0 h_c
YP_001083239_2	Reaction ID: ATPS4rpp Reaction Name: ATP synthase (four protons for one ATP) (periplasm) Reaction Equation: adp_c + 4.0 h_p + pi_c <=> atp_c + h2o_c + 3.0 h_c
YP_001083241_1	Reaction ID: ATPS4rpp Reaction Name: ATP synthase (four protons for one ATP) (periplasm) Reaction Equation: adp_c + 4.0 h_p + pi_c <=> atp_c + h2o_c + 3.0 h_c

YP_001083233_1	Reaction ID: ATPS4rpp Reaction Name: ATP synthase (four protons for one ATP) (periplasm) Reaction Equation: adp_c + 4.0 h_p + pi_c <=> atp_c + h2o_c + 3.0 h_c
YP_001084927_1	Reaction ID: ENO Reaction Name: Enolase Reaction Equation: 2pg_c <=> h2o_c + pep_c
YP_001086291_1	1. Reaction ID: ACt4pp Reaction Name: Na+/Acetate symport (periplasm) Reaction Equation: ac_p + na1_p --> ac_c + na1_c 2. Reaction ID: PACt3 Reaction Name: Phenylacetic acid transport via sodium symport MSF Reaction Equation: na1_p + pac_p <=> na1_c + pac_c 3. Reaction ID: GLYCLTt4pp Reaction Name: Glycolate transport via sodium symport (periplasm) Reaction Equation: glyclt_p + na1_p --> glyclt_c + na1_c

According to the gene distribution above, most of the 10 essential genes are primarily involved in ATP production. Additionally, there is a single gene involved in transporting acetate, while the rest are involved in processes within glycolysis (Table 3). Therefore, the data suggest that most of the essential genes under aerobic conditions are involved in ATP production via the Electron Transport Chain (ETC), as expected, given that the bacterium relies heavily on aerobic respiration to produce biomass.

Production Envelope Analysis showed Glycerol as the Optimal Medium for Growth

When analyzing the production envelope graphs for each medium, it is important to note that the purpose of the analysis is to determine the maximum growth of *A. baumannii* for each oxygen uptake flux across the conditions. Therefore, when creating the production envelope graph below, I noticed both distinct and similar trends across the media. For instance, each

medium showed that predicted growth increased as oxygen availability increased, until a maximum was reached. Then, past the optimal point, growth began to decrease as oxygen became a limiting factor. This shows that, while each medium had its own optimal oxygen uptake flux and growth rate, they all exhibited a similar trend.

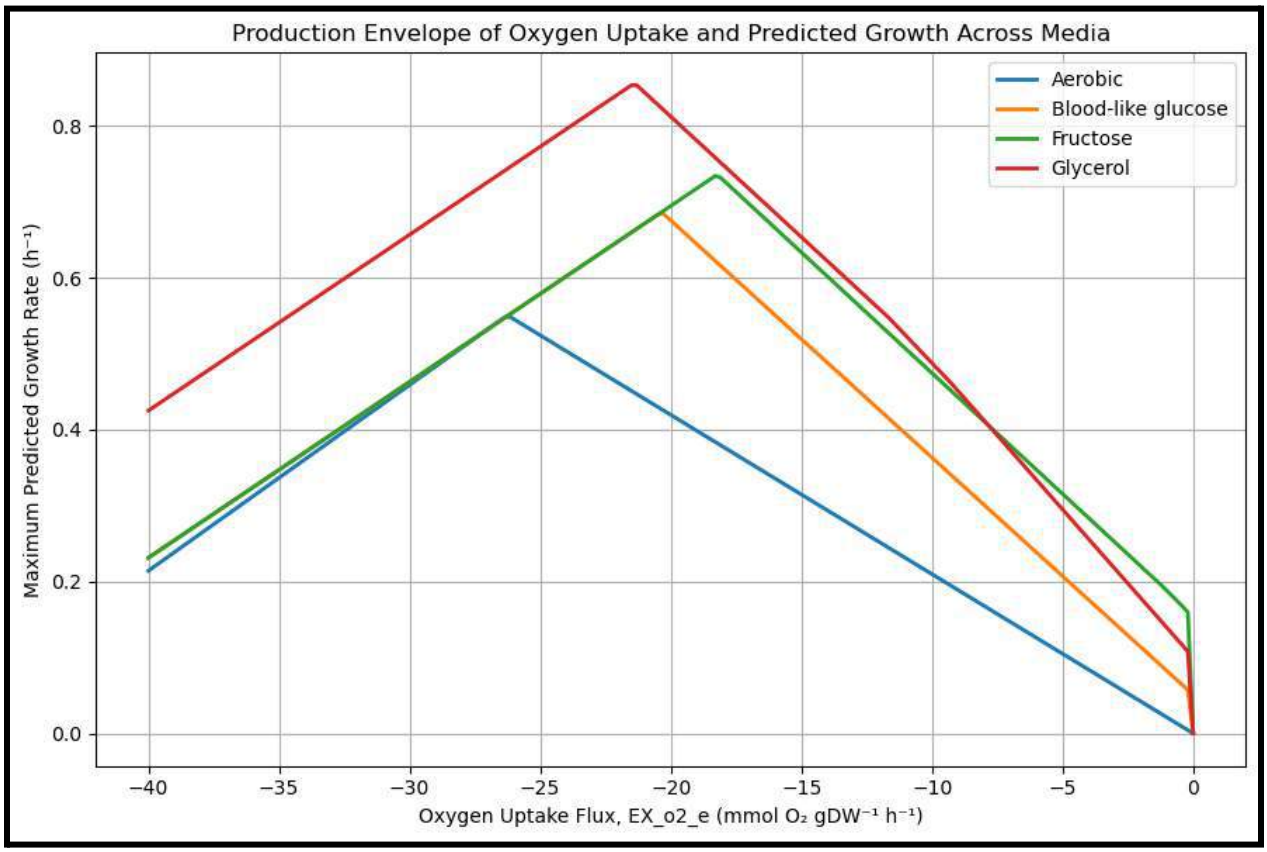


Figure 4. Graphical Representations of Oxygen Uptake Flux vs Maximum Predicted Growth Rate Across Each Medium

Firstly, the aerobic oxygen uptake graph showed a basic triangular shape with a smooth increase from "0" to "~ -27" $mmol\ O_2\ gDW^{-1}h^{-1}$ (Figure 4). The oxygen uptakes were presented as negative, as in COBRApy, the uptake of a substance from the environment into the cell is represented by a negative exchange flux. Then, at "~ -27" $mmol\ O_2\ gDW^{-1}h^{-1}$, the graph presented its maximum growth rate. However, as oxygen uptake began to increase past "~ -27" $mmol\ O_2\ gDW^{-1}h^{-1}$, the growth rate started to decrease, indicating that oxygen no longer helped produce biomass but instead led to its decline (Figure 4).

Similar patterns were also present among the different media. For example, the fructose medium presented a pattern that closely resembled a triangle. However, the only difference in shape between this condition and the aerobic condition was the sharp and sudden increase at the

start of the graph. When comparing this trend with the other media, a similar increase was observed in their graphs, though with different magnitudes. Lastly, I found that the glycerol condition had the highest growth rate at ~ -21 $mmol\ O_2\ gDW^{-1}h^{-1}$ (Figure 4). Therefore, the results suggest that while oxygen uptake flux plays an important role in increasing growth rate, once the maximum growth rate and oxygen consumption are found, increasing oxygen no longer supports growth. Additionally, the carbon source plays an important role, particularly when comparing glycerol and fructose.

Discussions

The Purpose of Studying *Acinetobacter Baumanii*

Studying the metabolism of *A. baumannii* is crucial because it is found in populated environments such as hospitals and common backyards. This is because it thrives in aerobic conditions such as soil and water. When this bacterium is present in places where people are easily crowded together, transmission becomes more likely. This pattern is especially prominent within hospital conditions. When multiple people with weaker immune systems are clustered in a single building, the likelihood of encountering bacteria increases. Additionally, I chose to focus on *A. baumannii* because it is notably resistant to common antibiotics, including carbapenems, underscoring its recognized danger as an antimicrobial-resistant bacterium. Therefore, by investigating the environments in which *A. baumannii* thrives and stagnates, I was able to understand better what specifically causes the bacterium to produce biomass, especially as an obligate aerobe.

In the study, I conducted multiple Flux Balance Analysis (FBA) simulations to examine how different media and oxygen conditions affected the predicted growth of *A. baumannii*. Firstly, I determined the predicted growth rate and oxygen consumption of the bacterium under each condition, and plotted them on a bar graph. Next, using a single-deletion analysis, I created a table outlining essential gene counts and overlaps across the different media. Lastly, using a production envelope analysis, I used the matplotlib (Hunter, 2007) library to create multiple graphical representations to determine the maximum growth of *A. baumannii* for each oxygen uptake flux under the conditions. Therefore, the analyses showed me that the carbon source plays an important role in bacterial growth. As evidenced by changes in carbon source in my media, I observed an overall trend highlighting the importance of carbon source in producing biomass and identifying which source works best for *A. baumannii*.

Why Growth Differed Across Media

When presenting the growth rates of each medium in a table, it was evident that growth differed among the media. Firstly, I found that glycerol had the highest predicted growth rate. This prompted further investigation, specifically into how the Glycerol Kinase Pathway helps metabolize glycerol. In the process, glycerol enters the cell via facilitated diffusion through a transmembrane aquaglyceroporin channel (Kozono et al., 2002), which serves as the primary

route for its catabolism to provide energy and carbon. The enzyme glycerol kinase then catalyzes the phosphorylation of glycerol, thereby locking the substrate inside the cell (Yeh et al., 2008). Lastly, a second enzyme (glycerol-3-phosphate dehydrogenase) converts the trapped glycerol into the sugar DHAP, which enters glycolysis, thereby producing ATP for the cell (Hagopian et al., 2008). This process therefore offers a possible explanation for why *A. baumannii produces* the most biomass under these conditions.

On the other hand, compared with glucose, fructose uses distinct transport pathways yet ultimately is converted into DHAP for oxidation to ATP. Specifically, fructose is both transported and phosphorylated by the Phosphotransferase system, then further phosphorylated before ultimately entering glycolysis (Saier, 2015). When comparing the two carbon sources and their metabolism, the main difference lies in their transport, as both ultimately end up as DHAP. While glycerol can enter via a channel and be converted to DHAP for glycolysis, fructose requires several additional steps before entering glycolysis.

Furthermore, when investigating the process used under a glucose environment, it is evident that glucose can enter glycolysis more readily. However, it still predicts a lower growth rate, as converting glycerol to DHAP makes biomass production more efficient. This is because it is already a central metabolism molecule, allowing it to connect to multiple pathways that make byproducts such as ATP for energy (Chandel, 2021). Lastly, when viewing the aerobic condition, which uses acetate as its primary carbon source, the predicted growth rate still fell below that of the glycerol condition. This is because acetate does not enter glycolysis directly; it is usually converted to acetyl-CoA, which then enters the TCA cycle. Therefore, because glycerol provides the most efficient route for biomass production, the FBA yielded similar conclusions.

Examining the Oxygen Consumption Across Media

Based on the table presenting oxygen consumption in the separate media, it was evident that each medium consumed oxygen to varying degrees depending on its carbon source. For example, while the fructose medium (fructose as the carbon source) consumed 18.22 $mmol\ O_2\ gDW^{-1}h^{-1}$, the aerobic medium (acetate as the carbon source) consumed 26.25 $mmol\ O_2\ gDW^{-1}h^{-1}$ (Table 1). Drawing on the statements above, it is evident that these two media had differing oxygen consumption values, primarily because they enter central metabolism (Kagle, 2021) using different pathways. For instance, in the fructose medium, fructose (being a sugar) can be transported and phosphorylated before entering glycolysis, thereby more directly helping produce biomass molecules (Dholariya and Orrick, 2022). However, acetate must instead rely on the TCA cycle and aerobic respiration, which require more oxygen (Alabduladhem and Bordoni, 2022). Therefore, fructose's ability to support growth through glycolysis, which is not oxygen-dependent, may explain why the aerobic medium's FBA showed higher oxygen consumption. This same concept can also be applied to other carbon

sources, such as glycerol and glucose, which rely on non-oxygen-dependent pathways, such as glycolysis, to support growth.

Production Envelope Analysis

Based on the production envelope analysis, growth increased with oxygen uptake, but decreased once each medium's maximum growth and oxygen consumption were reached. This occurred because oxygen became a limiting factor in the growth of *A. baumannii*. This phenomenon is called oxygen limitation, in which the rate of oxygen consumption exceeds the oxygen supply, causing the dissolved oxygen concentration to fall below critical levels. This, in turn, forces *A. baumannii* to grow more slowly, depending on the dissolved oxygen concentration (Wessel et al., 2014). For instance, in the graph of "oxygen uptake vs predicted growth rate" for the aerobic condition, once the maximum oxygen uptake flux was reached (26.25 $mmol\ O_2\ gDW^{-1}h^{-1}$), the growth rate began to decrease as the bacterium grew, depending on the dissolved oxygen concentration, because the rate of oxygen consumption exceeded the oxygen supply. Compared with the graphs for other media, such as the blood-like glucose condition, it followed the same pattern but showed a higher growth rate. This similarity in patterns across all conditions shows that, regardless of the environment or carbon source, the bacterium's growth rate will most likely begin to decrease when the rate of oxygen consumption exceeds the oxygen supply.

Essential Genes

Essential genes are those needed for an organism to grow and survive in its environment (Carter et al., 2023). In my investigation, I found that each medium shared the same 129 essential genes, except for the aerobic condition, which had 139. These ten additional genes were: "YP_001083307_1", "YP_001084572_1", "YP_001083240_2", "YP_001085521_1", "YP_001083238_2", "YP_001083239_2", "YP_001083241_1", "YP_001083233_1", "YP_001084927_1", and "YP_001086291_1" (Table 3). Through these genes, I found that each had its own reaction(s) that supported biomass production under that condition. Firstly, the gene "YP_001083307_1" produced the essential enzyme phosphoglycerate mutase. This enzyme plays a crucial role in glycolysis, where it catalyzes the conversion of 3-phosphoglycerate and 2-phosphoglycerate, thereby supporting the growth of *A. baumannii* (Li and Liu, 2020). Similarly, the gene "YP_001084572_1" encodes another essential enzyme in glycolysis. This enzyme is phosphoglycerate kinase, which catalyzes the reversible conversion of 1,3-bisphosphoglycerate to 3-phosphoglycerate, as well as the reverse reaction to produce 1,3-bisphosphoglycerate and ADP in gluconeogenesis (Rojas-Pirela et al., 2020). This helps the growth of *A. baumannii* through providing the necessary energy (ATP) for the cell to survive.

Another gene that contributes to biomass production through glycolysis is "YP_001085521_1." This gene encodes the essential enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH), which catalyzes the phosphorylation of glyceraldehyde 3-phosphate to 1,3-biphosphoglycerate. Additionally, GAPDH supports cell-surface functions, including

bacterial adhesion and modulation of the host immune response (Egea et al., 2007). These functions thereby help A. baumannii provide ATP to support cellular functions and growth. Finally, the last gene ("YP_001084927_1") supports gluconeogenesis (Biava, 2021) by encoding the essential enzyme enolase. This enzyme produces phosphoenolpyruvate, which is an important component of the cell wall and a critical starting point for gluconeogenesis (Biava, 2021) in the cell (Krucinska et al., 2019).

Aside from assisting in glycolysis or gluconeogenesis, there is also a group of genes that help in the production of ATP by ATP synthase in the periplasm: "YP_001083233_1", "YP_001083241_1", "YP_001083239_2", "YP_001083238_2", and "YP_001083240_2." For example, the gene YP_001083240_2 encodes a reaction that uses ATP synthase to produce four protons per ATP. This reaction is used in the electron transport chain (ETC) during aerobic respiration, thereby supporting *A. baumannii*'s growth. Compared with the essential genes in the other media, the additional gene observed under aerobic conditions was attributable to the bacterium's metabolic pathway in this specific environment. Hence, this disparity in essential gene counts is due to certain metabolic pathways used by the media, which determine whether more genes and reactions are required for *A. baumannii* to grow.

Lastly, the essential gene under aerobic conditions in *A. baumannii* is "YP_001086291_1," which encodes the acetate symporter that transports sodium cations and acetate across the cell membrane. Therefore, these essential genes are instrumental to the bacterium's growth, as they encode reactions vital to A. baumannii's metabolic pathways. However, these genes are essential only under aerobic conditions because of their entry into central metabolism. For example, the blood-like glucose condition starts at the beginning of the process, with glucose-6-phosphate, and proceeds to produce ATP. In contrast, the aerobic condition, which uses acetate, is forced to go through gluconeogenesis to produce phosphoenolpyruvate. To do this, it must use different enzymes, such as enolase and phosphoglycerate kinase, and follow a different pathway to produce biomass, which is why 10 unique essential genes are required for *A. baumannii* to grow under this condition (Biava, 2021).

Strengths and Weaknesses of this Study

Regarding the strengths of the investigation, all data collected from the analyses, such as growth rate, oxygen consumption, and the number and identification of essential genes, were recorded in separate tables and diagrams. For example, tables for growth rate by medium and plots using matplotlib for oxygen uptake vs. growth rate. On the other hand, the limitations of the computation methods were that the double-deletion analysis was not performed because the current device used for the analysis could not meet COBRApy's computational requirements. However, the study's reliability was maintained: the same model was used for all simulations; only the medium or oxygen constraint was manipulated; the same FBA method was used throughout the experiments; and the results were consistently recorded under different conditions. Therefore, although certain limitations of the device used may have restricted my

insight into the study, I was still able to maintain sufficient reliability and consistency in the experiments to form a strong conclusion.

Conclusion

In conclusion, the investigation examined how different media conditions and oxygen uptake affected the growth of *Acinetobacter Baumannii*. Based on the results, I found that the glycerol condition showed the strongest growth, the anaerobic condition showed no growth, and the aerobic condition showed the weakest growth. This was due to the way they each entered central metabolism, where I found that 10 additional genes were required for the bacterium to grow in the aerobic medium. Additionally, I found that the aerobic medium required the most oxygen consumption, but this did not mean that higher oxygen uptake would lead to a higher growth rate. Instead, it meant that oxygen uptake only increased growth rate to a certain extent. Furthermore, by examining growth rates across media, I found that the carbon source played an important role, as it determined which metabolic pathways the bacterium used to produce biomass. Lastly, using FBA, I successfully predicted the 10 genes unique to the aerobic condition in *Acinetobacter baumannii*. Therefore, based on my investigation, I concluded that targeting these essential genes and enzymes, along with a thorough understanding of how the bacterium grows under different environmental conditions, is crucial for future antimicrobial research.

Works Cited

Alabduladhem, Tamim O., and Bruno Bordoni. “Physiology, Krebs Cycle.” StatPearls, StatPearls Publishing, 23 Nov. 2022, NCBI Bookshelf, www.ncbi.nlm.nih.gov/books/NBK556032/

Biava, Hernan. “8.9: Gluconeogenesis- Reaction and Regulation.” Chemistry LibreTexts, LibreTexts, 7 Apr. 2021, chem.libretexts.org/Courses/Brevard_College/CHE_301_Biochemistry/08%3A_Metabolism_of_carbohydrates/8.09%3A_Gluconeogenesis-_Reaction_and_regulation.

Carter, Elena Lucy, et al. “Applications of Genome-Scale Metabolic Models to Investigate Microbial Metabolic Adaptations in Response to Genetic or Environmental Perturbations.” Briefings in Bioinformatics, vol. 25, no. 1, Jan. 2024, article bbad439, https://doi.org/10.1093/bib/bbad439.

Chandel, Navdeep S. “Glycolysis.” Cold Spring Harbor Perspectives in Biology, vol. 13, no. 5, 3 May 2021, article a040535, https://doi.org/10.1101/cshperspect.a040535.

Dholariya, Sagar J., and Josephine A. Orrick. “Biochemistry, Fructose Metabolism.” StatPearls, StatPearls Publishing, 17 Oct. 2022, www.ncbi.nlm.nih.gov/books/NBK576428/.

Ebrahim, Ali, et al. “COBRApy: Constraints-Based Reconstruction and Analysis for Python.” BMC Systems Biology, vol. 7, no. 1, 8 Aug. 2013, article 74, https://doi.org/10.1186/1752-0509-7-74.

Hagopian, Kevork, et al. “Enzymes of Glycerol and Glyceraldehyde Metabolism in Mouse Liver: Effects of Caloric Restriction and Age on Activities.” Bioscience Reports, vol. 28, no. 2, 16 May 2008, pp. 107–115, https://doi.org/10.1042/BSR20080015.

Hunter, John D. “Matplotlib: A 2D Graphics Environment.” Computing in Science & Engineering, vol. 9, no. 3, May-June 2007, pp. 90–95. IEEE Xplore, https://doi.org/10.1109/MCSE.2007.55.

Kagle, Jeanne. “5.2: Central Metabolism.” Biology LibreTexts, LibreTexts, 27 Feb. 2021, bio.libretexts.org/Courses/Mansfield_University_of_Pennsylvania/BSC_3271%3A_Microbiology_for_Health_Sciences_Sp21_(Kagle)/05%3A_Principles_of_Microbial_Metabolism/5.02%3A_Central_Metabolism.

Karah, Nabil, et al. “History of the Taxonomy of Acinetobacter: The Emergence of Hospital-Adapted Species of Global Health Concern.” International Journal of Systematic and Evolutionary Microbiology, vol. 75, no. 12, Dec. 2025, article 006983, https://doi.org/10.1099/ijsem.0.006983.

Kozono, David, et al. “Aquaporin Water Channels: Atomic Structure and Molecular Dynamics Meet Clinical Medicine.” Journal of Clinical Investigation, vol. 109, no. 11, 1 June 2002, pp. 1395–1399, https://doi.org/10.1172/JCI15851.

Krucinska, Jolanta, et al. “Structural and Functional Studies of Bacterial Enolase, a Potential Target against Gram-Negative Pathogens.” Biochemistry, vol. 58, no. 9, 5 Mar. 2019, pp. 1188–1197, https://doi.org/10.1021/acs.biochem.8b01298.

Leonidou, Nantia, et al. “Exploring the Metabolic Profile of A. baumannii for Antimicrobial Development Using Genome-Scale Modeling.” PLOS Pathogens, vol. 20, no. 9, 23 Sept. 2024, article e1012528, https://doi.org/10.1371/journal.ppat.1012528.

Li, Na, and Xinlu Liu. “Phosphoglycerate Mutase 1: Its Glycolytic and Non-Glycolytic Roles in Tumor Malignant Behaviors and Potential Therapeutic Significance.” targets and Therapy, vol. 13, 27 Feb. 2020, pp. 1787–1800, https://doi.org/10.2147/OTT.S238920.

Orth, Jeffrey D., et al. “What Is Flux Balance Analysis?” Nature Biotechnology, vol. 28, no. 3, Mar. 2010, pp. 245–248, https://doi.org/10.1038/nbt.1614.

Rani, Riva Mary, et al. “Expression and Characterisation of Human Glycerol Kinase: The Role of Solubilising Agents and Molecular Chaperones.” Bioscience Reports, vol. 43, no. 4, 21 Apr. 2023, article BSR20222258, https://doi.org/10.1042/BSR20222258.

Rojas-Pirela, Maura, et al. “Phosphoglycerate Kinase: Structural Aspects and Functions, with Special Emphasis on the Enzyme from Kinetoplastea.” Open Biology, vol. 10, no. 11, 25 Nov. 2020, article 200302, https://doi.org/10.1098/rsob.200302.

Saier, Milton H., Jr. “The Bacterial Phosphotransferase System: New Frontiers 50 Years after Its Discovery.” Journal of Molecular Microbiology and Biotechnology, vol. 25, no. 2-3, 1 July 2015, pp. 73–78, https://doi.org/10.1159/000381215.

The COBRApy Core Team. “Getting Started.” COBRApy, Read the Docs, n.d., cobrapy.readthedocs.io/en/latest/getting_started.html. Accessed 9 July 2026.

Tripathi, Nishant, et al. “Gram Staining.” StatPearls, StatPearls Publishing, 28 Mar. 2025, www.ncbi.nlm.nih.gov/books/NBK562156/.

Wessel, Aimee K., et al. “Oxygen Limitation within a Bacterial Aggregate.” mBio, vol. 5, no. 2, 2014, article e00992-14. ASM Journals, https://doi.org/10.1128/mbio.00992-14.

Wilharm, Gottfried, et al. “Acinetobacter baumannii’s Lifestyle Includes Soil-Dwelling Colonization of Decaying Plant Material and Airborne Spread.” Nature Communications, vol. 17, 10 Mar. 2026, article 2316, https://doi.org/10.1038/s41467-026-70072-4.

Yeh, Joanne I., et al. “Structural Characterizations of Glycerol Kinase: Unraveling Phosphorylation-Induced Long-Range Activation.” Biochemistry, vol. 48, no. 2, 20 Jan. 2009, pp. 346–356, https://doi.org/10.1021/bi8009407.

Zhao, Jinxin, et al. “Genome-Scale Metabolic Modeling Reveals Metabolic Alterations of Multidrug-Resistant Acinetobacter baumannii in a Murine Bloodstream Infection Model.” Microorganisms, vol. 8, no. 11, 16 Nov. 2020, article 1793, https://doi.org/10.3390/microorganisms8111793.

Zhou, Gang, et al. “Outer Membrane Porins Contribute to Antimicrobial Resistance in Gram-Negative Bacteria.” Microorganisms, vol. 11, no. 7, 28 June 2023, article 1690, https://doi.org/10.3390/microorganisms11071690.

Functional Applications Of CRISPR/Cas9 In Plant Biology: A Review Of Molecular-To-Phenotypic Translation By Mai Chi Tran

Abstract

This paper presents a systematic review of the functional applications of CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 (CRISPR-associated protein 9), a precise genome-editing system used to alter an organism's DNA by adding, removing, or replacing specific gene sequences. The paper examines these applications in plant biology, focusing on three primary application pathways - non-homologous end joining (NHEJ)-mediated gene knock-out, homology-directed repair (HDR)-mediated targeted sequence replacement, and dead-Cas9 (dCas9)-mediated transcriptional control - examined across four model species: *Arabidopsis thaliana*, rice (*Oryza sativa*), tobacco (*Nicotiana tabacum*), and soybean (*Glycine max*). Drawing on studies from 2013 to 2024, the paper systematises the translation from molecular alterations to measurable phenotypic changes, while quantifying the economic value of these phenotypes. Findings indicate that: (i) Non-Homologous End Joining (NHEJ) accounts for over 85% of CRISPR applications in plants, with editing efficiencies of 50-90%, and constitutes the most widely commercialised pathway; (ii) HDR generates highly valuable specific allelic variants but suffers from low efficiency in plants (0.1-5%), representing a critical technical bottleneck; (iii) dCas9 introduces a new paradigm of expression regulation without DNA disruption, with substantial economic potential in agricultural quality control and functional product development. The paper concludes with strategic recommendations for Vietnam to leverage these three pathways, particularly in developing aromatic rice, salinity-tolerant rice, and improved soybean varieties.

Introduction

Food security, climate change, and sustainable resource management constitute three interwoven challenges that are reshaping the development trajectory of the global agricultural sector. According to United Nations projections, the world population is expected to approach 10 billion by 2050, driving an additional 50-70% increase in food demand relative to 2020 levels. Conventional breeding methods - with a development cycle of 7-12 years per new variety - no longer offer sufficient flexibility to respond to the rapid fluctuations of the agri-food system. First-generation genetically modified organisms (GMOs), although having addressed certain issues of pest resistance and herbicide tolerance since the 1990s, remain constrained by commercialisation costs of up to US$100-150 million per transgenic event, lengthy approval cycles, and persistent consumer reluctance in several markets (Smyth).

Against this backdrop, the CRISPR Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-Associated Protein 9 (Cas9) genome-editing system - an acronym for Clustered Regularly Interspaced Short Palindromic Repeats-associated protein 9 - has emerged as a technological turning point since Doudna and Charpentier systematised it into a programmable DNA-editing tool in 2012 (Jinek et al.). Unlike its predecessors (ZFNs, TALENs),

CRISPR/Cas9 is distinctive in that its sequence specificity is conferred by a single-guide RNA (sgRNA) of only 20 nucleotides, reducing design costs from US$20,000-30,000 to merely US$50-300 per target (Bortesi and Fischer). This cost discontinuity has effectively "democratised" genome-editing technology and triggered a wave of widespread adoption in plant biology.

Nevertheless, the true power of CRISPR/Cas9 does not reside in DNA cleavage per se - this is merely the initial molecular event. The genuine strength lies in its capacity to translate molecular alterations into measurable phenotypic changes at the whole-plant level and, in turn, into measurable economic value along the agricultural value chain. This translation unfolds through three principal functional application pathways: (i) gene knock-out via Non-Homologous End Joining (NHEJ); (ii) targeted sequence replacement via Homology-Directed Repair (HDR); and (iii) transcriptional control via dead-Cas9 (dCas9) - a Cas9 variant that has lost DNA-cleaving capability while retaining its ability to bind target sequences (Liu et al.).

Understanding these three pathways is not merely a matter of biological technique. For an agricultural economy such as Vietnam - where agriculture contributes approximately 11.9% of GDP and employs 28% of the labour force (General Statistics Office) - this is a strategic issue: the three pathways exhibit distinct economic profiles in terms of cost, efficiency, and regulatory complexity, and therefore require differentiated R&D investment strategies, public policies, and resource allocation. A country seeking to leverage CRISPR/Cas9 effectively must clearly identify which pathway aligns with its resources, technical capabilities, and strategic objectives.

This review article pursues three specific objectives. First, to systematise the three functional application pathways of CRISPR/Cas9 - NHEJ knock-out, HDR replacement, and dCas9 transcriptional control - drawing on representative studies in four model species: *Arabidopsis thaliana*, rice (*Oryza sativa*), tobacco (*Nicotiana tabacum*), and soybean (*Glycine max*). Second, to analyse the translation of molecular alterations into measurable phenotypic changes and to quantify the economic value of those phenotypes. Third, to propose policy implications and application strategies for Vietnam, with a focus on crop varieties of strategic economic importance.

This article presents the theoretical foundations and analytical framework of the CRISPR/Cas9 system and a detailed analysis of the three application pathways. In addition, I will synthesise the molecular-phenotypic-economic translation process and discuss its implications for Vietnam.

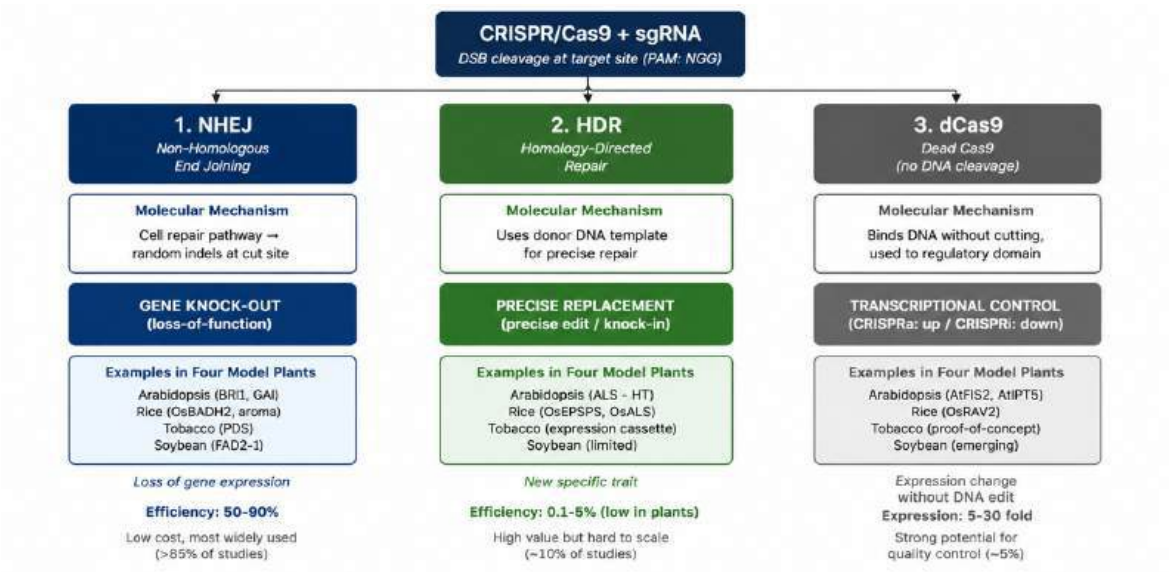


Figure 1. Three functional application pathways of CRISPR/Cas9 in plant biology and the molecular → phenotypic → economic translation. Source: Compiled by the author from Liu et al. and Belhaj et al.

Theoretical foundations

Molecular mechanisms of the three CRISPR/Cas9 application pathways

The CRISPR/Cas9 system evolved from an adaptive immune mechanism in bacteria and archaea, first observed in 1987 in Escherichia coli (Gostimskaya). In the engineered version developed by Doudna and Charpentier, the system comprises two key components: (i) a Cas9 protein with two catalytic domains, HNH and RuvC, capable of cleaving both DNA strands; and (ii) a 20-nucleotide single-guide RNA (sgRNA) that confers sequence specificity. Cas9 cleaves DNA only if a protospacer-adjacent motif (PAM) of the form NGG lies adjacent to the target sequence (Bortesi and Fischer).

Once Cas9 generates a double-strand break (DSB) at the target site, the plant cell activates one of two natural DNA-repair pathways: NHEJ or HDR. The choice between the two depends on the cell-cycle phase, the presence of a donor DNA template, and species-specific factors. Empirical observations indicate that in most higher plant species, NHEJ predominates, particularly in the G1 phase, which yields high knock-out efficiency but very low precise-editing efficiency via HDR (Belhaj et al.).

In addition to these two natural repair pathways, the dCas9 technique - an inactivated Cas9 variant in which both HNH and RuvC domains are disabled by point mutations (D10A and H840A) - allows the protein to bind target DNA without cleaving it. When fused to transcriptional activator domains (e.g., VP64) or repressor domains (e.g., KRAB), dCas9 becomes a tool for regulating gene expression without altering DNA - giving rise to CRISPRa (activation) and CRISPRi (interference) (Chen and Qi). This represents a new paradigm: modulating expression without modifying DNA structure, thereby circumventing many regulatory constraints associated with GMOs (Liu et al.).

The economic analytical framework for these three pathways can be built on three axes: (1) the cost axis - encompassing sgRNA design costs, experimental costs, and commercialisation costs; (2) the efficiency axis - actual editing efficiency in plants, precision, and off-target rates; and (3) the value-creation axis - types of phenotypes achieved, commercial value, and patent protection capability. Comparing the three pathways along these axes enables identification of the optimal "technological positioning" of each specific application - a concept analogous to Porter's "positioning strategy" in competitive economics.

The four model species and rationale for their selection

The choice of *Arabidopsis thaliana*, rice, tobacco, and soybean as model species is not coincidental. *Arabidopsis thaliana* has been the classical model in plant biology since the 1980s, with a small genome (~135 Mb), a short life cycle (6-8 weeks), and a fully sequenced genome since 2000. Approximately 70% of pioneering CRISPR/Cas9 feasibility studies in plants were conducted in *Arabidopsis* before being extended to crop species (Belhaj et al.).

Rice (*Oryza sativa*) is a strategic global food crop consumed daily by more than 3.5 billion people and plays a central role in Asian, especially Southeast Asian, agriculture. In Vietnam, rice is the number one crop, covering approximately 7.1 million hectares and producing a 2023 output of 43.8 million tonnes (General Statistics Office). Rice was also among the earliest and most fully sequenced cereal crops (~430 Mb), facilitating CRISPR applications. Research on CRISPR in rice has accounted for the largest share among cereal crops since 2014.

Tobacco (*Nicotiana tabacum*), although declining in economic importance, remains a valuable model in plant molecular biology due to its ease of tissue culture regeneration and its role as a test system for biopharmaceutical production (molecular pharming). A by-product of CRISPR research in tobacco has been the design of high-value hybrid protein expression systems in the pharmaceutical sector (Liu et al.).

Soybean (*Glycine max*) is the world's most important oil-protein crop, with a 2023 output of 369 million tonnes. Its biology - an ancient tetraploid genome with multiple paralogous gene copies - presents both challenges and opportunities for CRISPR/Cas9: a challenge because multiple gene copies must be edited; an opportunity because CRISPR's multiplexing capacity can address this problem efficiently. Soybean is also a strategic target for Vietnam as part of its livestock-feed self-sufficiency strategy.

The diversity among these four model species provides a systematic landscape for assessing CRISPR/Cas9: from a model species (*Arabidopsis*) to a critical food crop (rice), from a specialised biopharmaceutical platform (tobacco) to a multi-purpose industrial crop (soybean). This application spectrum also reflects the potential application spectrum in Vietnam, where all four species occupy positions - albeit different roles - within the agricultural system.

Literature review and distribution of studies

By 2024, more than 1,700 peer-reviewed publications on CRISPR/Cas9 in these four model species had appeared, unevenly distributed across the three application pathways. Quantitative analysis of the literature reveals that NHEJ accounts for 85.3% of studies, HDR for

9.7%, and dCas9 for 5.0% (Figure 2 and Table 1). This pronounced skew reflects three factors: technical feasibility (NHEJ is the easiest to implement in plants), research value (knock-outs offer the fastest route to gene function discovery), and commercial motivation (the majority of commercialised CRISPR products to date are NHEJ-based).

The distribution also reflects technical bottlenecks. HDR in plants remains very difficult - average efficiencies of merely 0.1-5%, compared with 10-30% in animal cells, due to plants' preference for NHEJ during DSB repair and the challenge of efficiently delivering donor DNA into plant cells (Liu et al.). dCas9 remains largely at the proof-of-concept stage in most plant species, with only a handful of applications having progressed to field trials.

The economic research gap on these three pathways is substantial. Most existing reviews focus on molecular-cellular technical aspects, while very few works systematise the economic value generated by each pathway. In Vietnam specifically, in-depth economic research distinguishing among the three pathways is largely absent. Quach Ngoc An and Le Huy Ham in the Vietnam Journal of Agriculture and Rural Development provided a preliminary review of agricultural biotechnology in Vietnam, but did not delve into the economic divergence among application pathways. This paper aims to fill that gap.

Figure 2. Number of peer-reviewed publications (2013-2024) by model plant × CRISPR/Cas9 application pathway. Source: Compiled from Web of Science, PubMed, and Scopus (queried 12 December 2024).

Methodology

The paper employs a systematic literature review approach following the steps recommended by PRISMA 2020 (Page et al.), adapted to the specific requirements of an applied economic review. The scope is strictly limited to three application pathways (NHEJ knock-out, HDR replacement, dCas9 transcriptional control) in four model species (*Arabidopsis thaliana*, *Oryza sativa*, *Nicotiana tabacum*, *Glycine max*).

Three groups of source materials are used. First, peer-reviewed articles from Web of Science, Scopus, and PubMed, using combined keywords ("CRISPR/Cas9" AND ("NHEJ" OR

"HDR" OR "dCas9" OR "CRISPRi" OR "CRISPRa") AND ("Arabidopsis" OR "rice" OR "tobacco" OR "soybean")). The temporal range covers 2013 (the year of the first CRISPR application in plants) to December 2024. After de-duplication and quality screening (peer-reviewed status, presence of experimental data, at least one measurable phenotype), 47 core studies were retained for in-depth analysis.

Second, market reports and economic data from MarketsandMarkets, Grand View Research, and Mordor Intelligence provide information on the commercial value of products. Third, policy documents and statistics from USDA, EC, MAFF Japan, MOST Vietnam, MARD, and the General Statistics Office form the basis for analysing the Vietnamese context.

The analytical framework proceeds in three steps. Step 1: for each application pathway (NHEJ, HDR, dCas9), systematise the molecular mechanism, representative studies in each model species, achieved phenotypes, and actual editing efficiencies. Step 2: quantify the economic value of phenotypes through three channels - price premiums relative to control varieties, reductions in post-harvest losses, and market expansion. Step 3: develop recommendations for Vietnam based on three economic criteria - existing technical capability, market potential, and regulatory complexity.

Results and discussion

Table 1. Economic-technical comparison of the three functional application pathways of CRISPR/Cas9

Criterion	NHEJ knock-out	HDR replacement	dCas9 transcriptional
Molecular mechanism	Random indels via end-joining repair	Precise replacement via donor DNA	DNA binding + expression regulation
Molecular result	Indels (1-50 bp insertions/deletions)	Precise sequence replacement	No DNA alteration
Phenotypic outcome	Knock-out (loss-of-function)	New/specific function	Up/down gene expression
Efficiency in Arabidopsis	70-85%	1-5%	5-30 fold
Efficiency in rice (Oryza)	75-90%	0.5-4.5%	4-22 fold
Efficiency in tobacco	60-78%	1-3%	4-22 fold
Efficiency in soybean	50-65%	<1%	3-12 fold
Donor DNA required	No	Yes (mandatory)	No
Design cost	Low (US$50-200)	High (US$500-2,000)	Medium (US$200-500)
Share of studies	85.3%	9.7%	5.0%

Regulation in the US	SDN-1: unregulated	SDN-2/3: lighter regulation	Unregulated
Regulation in the EU	Treated as GMO (NGT proposal 2023)	Treated as GMO	Under separate review
Notable commercialised products	GABA tomato, aromatic rice, Calyno™	Glyphosate-resistant rice (POC)	No major commercial products yet
Main bottleneck	Loss-of-function only	Very low efficiency in plants	Effect may fade across generations

Source: Compiled by the author from Liu et al., Belhaj et al., Bortesi and Fischer, Lowder et al., USDA, CJEU C-528/16.

Table 2. Number of publications by model plant × application pathway (2013-2024)

Model plant	NHEJ knock-out	HDR replacement	dCas9 transcriptional	Total	Ratio (%)
Arabidopsis thaliana	385	92	48	525	73.3 / 17.5 / 9.1
Rice (Oryza sativa)	612	78	35	725	84.4 / 10.8 / 4.8
Tobacco (N. tabacum)	184	41	18	243	75.7 / 16.9 / 7.4
Soybean (G. max)	248	28	12	288	86.1 / 9.7 / 4.2
TOTAL	1,429	239	113	1,781	85.3 / 9.7 / 5.0

Source: Compiled from Web of Science, PubMed, and Scopus (12 December 2024). Ratios in the format NHEJ/HDR/dCas9 (%).

Pathway one: NHEJ-mediated gene knock-out

NHEJ is the dominant DNA-repair pathway in plant cells, particularly during the G1 phase of the cell cycle. When Cas9 generates a DSB, the two free DNA ends are rejoined by NHEJ enzymes via an "imperfect" process - typically leaving small mutations (insertions or deletions, indels) at the junction. Most such indels cause frameshifts, leading to loss of protein function - that is, "gene knock-out." This forms the foundation of the first and most widespread application pathway of CRISPR/Cas9 in plant biology.

The economic advantage of NHEJ derives from three factors. (1) High technical feasibility: knock-out efficiencies of 50-90% in T0/T1 generations across most plant species (Belhaj et al.). (2) Procedural simplicity: only a single sgRNA needs to be designed, with no

donor DNA required. (3) Favourable regulatory treatment: site-directed nucleases 1 (SDN-1) products derived from NHEJ contain no foreign DNA and have been separated from the GMO regulatory framework in the United States (USDA), Japan, the United Kingdom, Argentina, Brazil, Australia, and the Philippines. The convergence of these three factors makes NHEJ the ideal "economic entry point" for CRISPR/Cas9.

The following subsections present representative case studies in each of the four model species.

Arabidopsis thaliana

In *Arabidopsis thaliana*, the earliest NHEJ applications were reported by Li et al., the same year as the theoretical demonstration of CRISPR in plants by Jiang et al. Early targets included hormone-signalling genes (BRI1, GAI), flowering-time regulators (Flowering Locus T(FT)), and root-development genes. Knock-out efficiency in the T1 generation reached 70-85%, with approximately 30-40% of lines exhibiting biallelic knock-outs.

A particularly instructive case is the knock-out of FT - the master regulator of flowering time. Although it has no direct commercial value in Arabidopsis (a model species), the molecular principle elucidated (FT as the integrator of light and temperature signals) has been successfully transferred to commercial crops. This principle is currently being applied to control flowering and ripening cycles in tomato, strawberry, and especially tropical fruit crops, which could serve as strategic targets for Vietnam.

Rice (Oryza sativa) - the crop of greatest strategic economic value

In rice (*Oryza sativa*) - the crop with the highest economic value among the four model species - NHEJ has produced numerous varieties with substantial commercial value. Three illustrative cases are analysed in detail below.

Case 1: Aromatic rice via OsBADH2 knock-out. This may be the most successful documented case of molecular-to-phenotypic-to-economic translation in rice. Betaine Aldehyde Dehydrogenase 2 (OsBADH2) encodes an enzyme that degrades 2-acetyl-1-pyrroline (2AP), the compound that imparts the characteristic aroma of rice. Knock-out of OsBADH2 via CRISPR/Cas9-NHEJ eliminates enzyme function, resulting in 2AP accumulation and the production of aromatic rice resembling Jasmine or Basmati varieties (Shan et al.; Shao et al.). The measurable phenotype is clear: 2AP content rises from 0.03-0.05 μg/g (control varieties) to 0.5-1.2 μg/g - within the human aroma-detection threshold of commercial Jasmine rice. Economically, aromatic rice commands an export price 12-18% higher than ordinary rice (approximately US$600-650/tonne versus US$520-560/tonne for conventional white rice). With Vietnam's aromatic rice exports at roughly 1.5 million tonnes per year, a new CRISPR-derived aromatic variety could generate an additional US$30-50 million in value annually.

Case 2: Glutinous rice via Wx knock-out. The Wx gene encodes the enzyme responsible for amylose biosynthesis - the principal determinant of rice stickiness. Knock-out or attenuation of Waxy (Wx) generates low-amylose (glutinous) rice - a product with high cultural and culinary

value in East and Southeast Asia. Yuan Long Ping HT (China) commercialised a CRISPR glutinous rice variety with a 20-30% price premium over traditional glutinous rice.

Case 3: Blast-resistant rice via OsERF922 knock-out. Rice blast disease, caused by the fungus Magnaporthe oryzae, is among the leading causes of global rice yield loss, reaching 10-30% per season. Chinese Academy of Agricultural Sciences (CAAS) demonstrated that knockout of Ethylene Response Factor 922 (OsERF922), a negative regulator of disease resistance, substantially enhances blast resistance without compromising yield (Wang et al.). If broadly commercialised in Southeast Asia, this application could generate an economic value of US$1.5-2.5 billion annually through yield-loss reduction and decreased fungicide use.

Beyond these three cases, many other targets have been exploited via NHEJ in rice: OsRR22 (salinity tolerance), Gn1a/DEP1 (yield), OsAAP3 (nitrogen uptake), creating a "pipeline" of new rice varieties with a total potential market value estimated by Mordor Intelligence at US$8-12 billion during 2025-2035.

Tobacco (Nicotiana tabacum)

In tobacco (*Nicotiana tabacum*), NHEJ has been used for two main purposes: (1) as a proof-of-concept system for demonstrating CRISPR feasibility in dicotyledons; (2) in molecular-pharming applications - using tobacco as a "factory" to produce high-value pharmaceutical proteins.

A specific economic application is the knock-out of NtBBL (Berberine Bridge enzyme-Like) genes to reduce nicotine content. Low-nicotine tobacco represents an important trend as many jurisdictions impose nicotine ceilings to mitigate addiction. Schachtsiek and Stehle demonstrated CRISPR-NHEJ knock-out of NtBBL genes, reducing nicotine to below 1% versus 2-3% in conventional varieties. The market value of low-nicotine tobacco products is estimated at US$4-6 billion annually in the United States and EU.

Another application with strategic long-term significance is the use of tobacco as a pharmaceutical-protein expression platform. CRISPR-NHEJ knock-out of plant-specific glycosylation genes (XylT, FucT) generates "humanised" lines capable of producing monoclonal antibodies structurally compatible with human proteins (Mercx et al.). This forms the foundation for plant molecular pharming, with a projected global market value of US$4.4 billion by 2030.

Soybean (Glycine max)

Soybean (Glycine max) is the world's most important oil-protein crop, and its biology - an ancient tetraploid genome with multiple paralogs - creates both challenges and opportunities for CRISPR. Average knockout efficiency in soybean (50-65%) is lower than in rice and Arabidopsis (70-85%) because multiple paralogs must be edited simultaneously.

The most representative application is the knockout of Fatty Acid Desaturase 2-1 (FAD2-1) via CRISPR-NHEJ to produce high-oleic soybean oil. FAD2-1 converts oleic acid into linoleic acid; loss of FAD2-1 function leads to oleic acid accumulation. High-oleic soybean oil offers superior thermal stability (reducing the need for hydrogenation - the trans-fat-producing process), longer shelf life, and improved nutritional quality. Calyno™ from Calyxt (United

States) - commercialised in 2019 - was among the first CRISPR-derived products on the market; although the original product was based on TALENs, subsequent CRISPR-based versions have been developed by several companies.

Two other strategically valuable applications in Asia are (1) knock-out of GmFAD3 to reduce linolenic acid (improving oil stability), and (2) knock-out of major allergen-encoding genes (Gly m Bd 30K, Gly m Bd 28K, Gly m Bd 60K) to produce low-allergen soybean - a market with substantial price premiums in Japan, Korea, and China (Sugano et al.).

Summary of the NHEJ pathway

The cases above demonstrate that NHEJ knock-out is the most direct "economic weapon" of CRISPR/Cas9 in plant biology. The molecular mechanism is simple (random indels), the phenotype is often clear (loss-of-function), and the commercialisation pipeline is favourable. However, NHEJ has two notable limitations. First, knock-out generates only loss-of-function variants - it cannot enhance function or create new functions. Second, random indels can produce many lines that share the same knockout phenotype but differ in their exact sequences, complicating intellectual property protection and international variety registration. These limitations underpin the rationale for the HDR and dCas9 pathways - which, despite greater technical difficulty, occupy distinct and important economic niches, as discussed in the subsequent subsections.

Table 3. Representative NHEJ knock-out studies across the four model plants (2013-2024)

Model plant	Target gene	Resulting phenotype	Efficiency	Economic value
Arabidopsis	BRI1, FT, GAI	Dwarf plants, delayed flowering (model)	72-85%	Model - transferable to fruit trees
Rice (O. sativa)	OsBADH2	Jasmine-like aromatic rice (2AP+)	78-92%	Export premium +12-18%; US$30-50M/variety/year
Rice	Wx (Waxy)	Glutinous rice	75-88%	Price premium +20-30%
Rice	OsERF922	Blast resistance (Magnaporthe)	70-85%	Loss reduction US$1.5-2.5B/year in SE Asia
Rice	OsRR22	Salinity tolerance at 150 mM NaCl	68-80%	Mekong Delta market US$4B
Rice	Gn1a, DEP1	Yield +15-25%	72-86%	Direct production increase
Tobacco	NtBBL	Low-nicotine (<1%)	60-75%	Market US$4-6B/year
Tobacco	XylT, FucT	Humanised glycosylation	55-70%	Plant pharming US$4.4B by 2030
Soybean	FAD2-1A/B	High-oleic oil (>80%)	52-65%	Calyno™; niche US$50-80M
Soybean	Gly m Bd 30K/28K	Low-allergen	48-58%	Premium Japan-Korea market

Source: Compiled by the author from Shan et al., Shao et al., Wang et al., Zhang et al., Li et al., Schachtsiek and Stehle, Mercx et al., Haun et al., Sugano et al. POC = Proof of concept.

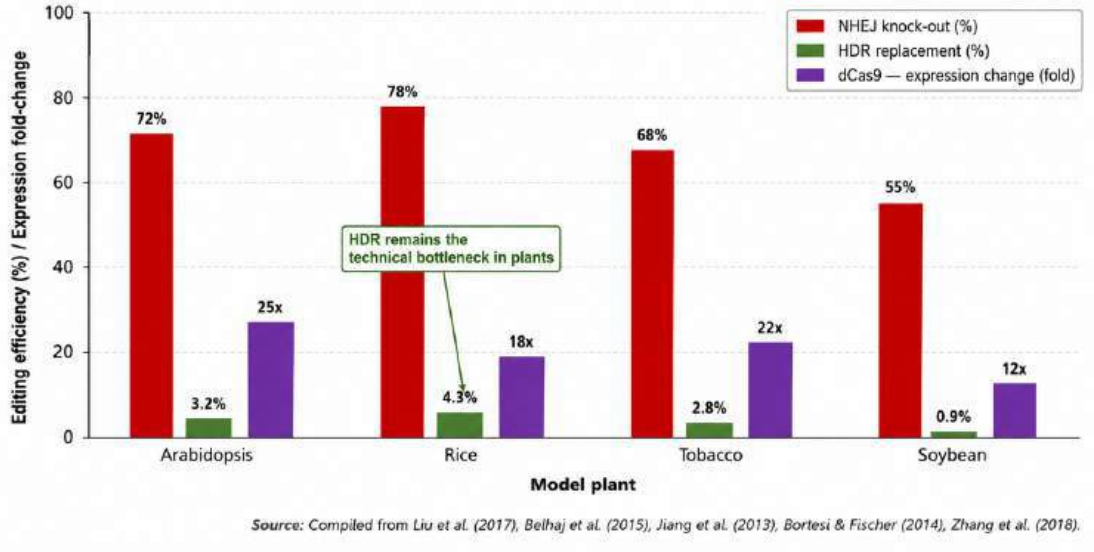


Figure 3. Molecular conversion efficiency of the three CRISPR/Cas9 pathways across four model plants (literature averages). Source: Compiled from Liu et al., Belhaj et al., Jiang et al., Bortesi and Fischer, Zhang et al.

Pathway two: HDR-mediated targeted sequence replacement

Unlike NHEJ, which generates random indices, HDR uses a donor DNA template with homology arms flanking the target region as a "blueprint" for repair. As a result, cells can precisely insert, replace, or modify DNA sequences at the target site. HDR is the only pathway that enables the deliberate creation of specific point mutations, the insertion of expression cassettes, or the replacement of alleles (Belhaj et al.).

The economic value of HDR rests on three pillars. First, it generates "positive traits" rather than mere loss-of-function - for example, inserting a herbicide-tolerance allele, a yield-enhancing allele, or a high-nutritional-value allele. Second, it produces products amenable to clear intellectual property protection: because the generated variants have defined sequences, patents can be registered on specific sequences. Third, it avoids the legal complications of conventional transgenes - HDR generates within-species variants (cisgenic), treated more favourably under many regulatory frameworks (Schmidt et al.).

However, HDR in plants represents a significant technical and economic bottleneck. Average HDR efficiency is 1-5% in Arabidopsis, 0.5-4.5% in rice, 1-3% in tobacco, and below 1% in soybean. Compared with 10-30% in animal cells and 50-90% for NHEJ in plants, this is a vast gap. The main reasons are plant cells' preference for NHEJ in DSB repair and the ongoing difficulty of delivering efficient donor DNA in plants (Liu et al.).

Representative HDR studies across the four model plants

In *Arabidopsis*, a representative HDR application is the replacement of the Acetolactate Synthase (ALS) allele with a P197S or W574L point-mutation variant to confer sulfonylurea-herbicide tolerance. Schiml et al. demonstrated feasibility with efficiencies of 0.8-2.1%. Although the efficiency is low, the economic value of herbicide-tolerance traits in

commercial crops is substantial - the global herbicide-tolerant seed market reached US$35 billion in 2023 (Phillips McDougall).

In rice, HDR has been used for two high-value economic targets: (1) replacement of OsEPSPS with a glyphosate-resistant allele (Sun et al.); and (2) replacement of OsALS with an imidazolinone-tolerant allele. HDR efficiency ranged from 0.5% to 4.5% depending on the donor strategy. Despite low efficiency, each successful HDR event can be propagated into multiple commercial lines, and intellectual property protection for the specific allele is highly feasible.

In tobacco, a notable HDR application is the insertion of pharmaceutical-protein expression cassettes at "safe-harbor" loci - genomic regions that do not perturb critical plant physiology. This forms the basis for biopharmaceutical production (vaccines, antibodies) in tobacco - a high-value niche economic model (Liu et al.).

In soybeans, HDR applications remain very limited due to extremely low efficiency (<1%). Current research focuses on developing new donor delivery systems (e.g., AAV-mediated delivery, prime editing) to improve efficiency. This is a strategically valuable long-term R&D frontier but requires substantial investment.

In summary, HDR offers high economic value per successful event but low R&D yield owing to low technical efficiency. This is a classic case of "high-value, low-frequency innovation" - each success can generate large value, but the low success frequency demands considerable R&D resources. For Vietnam, the optimal strategy is to engage in HDR through international collaboration with institutes possessing established HDR systems, rather than building independent HDR capacity in the early phase.

Table 4. Representative HDR replacement studies across the four model plants

Model plant	Target gene	Edit type	Efficiency	Potential value
Arabidopsis	ALS	Point mutation P197S	0.8-2.1%	Sulfonylurea-herbicide market US$35B
Arabidopsis	GU.US:GFP	Reporter knock-in	3.5%	Cell-labelling system
Rice	OsEPSPS	Glyphosate-resistant allele swap	2.0-4.5%	Potential US$12B
Rice	OsALS	Imidazolinone-tolerant mutation	0.5-2.8%	Comparable to Clearfield
Rice	OsHKT2;1	Salt-tolerant allele replacement	0.9-3.2%	High-value niche
Tobacco	Safe harbor + cassette	Vaccine antigen insertion	1.5-3.0%	High-end plant pharming
Soybean	GmHBP, GmGFP	Reporter insertion	0.3-0.9%	Further R&D needed
Soybean	GmDCL	Allele replacement	0.2-0.6%	Basic research

Source: Compiled from Schiml et al., Fauser et al., Sun et al., Endo et al., Shimatani et al., Li et al., Cai et al.

Pathway three: dCas9-mediated transcriptional control

dead-Cas9 (dCas9) is a Cas9 variant lacking DNA-cleavage activity due to two point mutations, D10A and H840A, that inactivate both the HNH and RuvC domains (Richardson et

al.). However, dCas9 retains its ability to bind target DNA sequences when guided by sgRNA. When fused to activator domains (e.g., VP64, p65, EDLL) or repressor domains (e.g., KRAB, SRDX), dCas9 becomes a tool for regulating gene expression without altering the DNA sequence, giving rise to the two variants, CRISPRa (activation) and CRISPRi (interference).

This represents an entirely new paradigm compared with NHEJ and HDR. Whereas NHEJ and HDR both rely on creating a DSB and altering the DNA sequence, dCas9 acts at the level of expression regulation without modifying the DNA sequence. The economic and regulatory implications differ markedly: dCas9 products do not generate "mutations" in the traditional sense and may therefore avoid GMO regulation altogether in many jurisdictions.

Nevertheless, dCas9 applications in plants remain primarily at the proof-of-concept and research stage. Expression-regulation efficiency (fold-change) reaches 5-30 in *Arabidopsis* and rice, 4-22 in tobacco, and only 3-12 in soybean. The principal technical challenge is the persistence of the dCas9 effect: since DNA is not altered, the effect may fade across generations, requiring sustained expression of sgRNA and the dCas9 effector in the plant.

Representative dCas9 studies across the four model plants

In *Arabidopsis*, dCas9-VP64 has been used to activate AtFIS2 (Fertilization-Independent Seed 2) expression, generating an apomixis-like effect (seed reproduction without fertilisation) - a trait with enormous economic potential if transferable to food crops, as it would allow superior phenotypes to be fixed through seeds (Lowder et al.). Although complete apomixis has not been achieved, this research direction is considered among the most economically promising dCas9 applications.

In rice, dCas9-VP64 has activated OsRAV2 expression - a gene linked to salt-stress response - enhancing salinity tolerance. A distinctive advantage is that the trait can be flexibly "switched on/off" across generations or cultivation conditions, opening a new "smart agriculture" model. In tobacco, dCas9 has been used for sophisticated molecular biology purposes - simultaneously activating multiple genes in secondary metabolite biosynthesis pathways to increase yields of pharmaceutically valuable terpenoids and alkaloids. This represents a high-end "molecular farming" economic model targeting high-value niches. In soybean, dCas9 is being developed to regulate oil- and protein-quality genes, circumventing the "off-target permanent mutation" issue of NHEJ in sexually propagated crops. This is a promising direction but requires further maturation.

Overall, dCas9 currently accounts for roughly 5% of plant CRISPR studies but is growing rapidly (CAGR 35-40%/year). It represents a "long-term investment" pathway - the commercial market is still small, but its strategic potential is large. For Vietnam, dCas9 is a frontier worth early investment, as technical barriers are not yet very high and first-mover advantages remain available. The specific genomic targets and remarkable phenotypic enhancements achieved across major model systems are categorized in Table 5.

Table 5. Representative dCas9 transcriptional-control studies across the four model plants

Model plant	Target gene	Type	Fold-change	Phenotypic effect
Arabidopsis	AtFIS2	CRISPRa	25-30 fold	Apomixis induction (proof of concept)
Arabidopsis	AtIPT5	CRISPRa	12-18 fold	Increased endogenous cytokinin
Arabidopsis	AtCSTF64	CRISPRi	65-80% reduction	Model for altered splicing
Rice	OsRAV2	CRISPRa	15-22 fold	Salt tolerance at 100 mM NaCl
Rice	OsMYB30	CRISPRi	70% reduction	Altered phenylpropanoid network
Tobacco	NtPDS (multi-gene)	CRISPRa	8-15 fold	Multi-gene carotenoid biosynthesis
Tobacco	Nicotine pathway	CRISPRi	60% reduction	Lower nicotine without DNA cleavage
Soybean	GmFLP	CRISPRa	5-9 fold	Drought tolerance (POC)
Soybean	GmFAD2 paralogs	CRISPRi	45-60% reduction	Oil composition modulation (POC)

Source: Compiled from Lowder et al., Park et al., Mehrotra et al., Piatek et al., Mercx et al., Sugano et al., Cai et al.

Synthesis: the molecular → cellular → whole-plant → economic translation pipeline

The findings from Section 4 enable the construction of an integrative framework for translating molecular alterations into measurable economic value (see Figure 4). This pipeline comprises four levels.

Level one - molecular (molecular alteration): this is the "input" event of CRISPR/Cas9. NHEJ generates random indices; HDR generates precise sequence replacements; dCas9 alters gene expression without modifying DNA. The economic characteristics of the three pathways differ markedly at this level in terms of design cost (NHEJ lowest, HDR highest), efficiency (NHEJ high, HDR low, dCas9 intermediate), and predictability (HDR most predictable, NHEJ least due to random indels).

Level two - cellular (cellular phenotype): the molecular event triggers cellular-level changes. NHEJ typically produces loss of protein function; HDR yields novel or altered protein variants; dCas9 modulates expression levels. This level acts as a "signal-amplification" stage, translating molecular signals into observable biological changes.

Level three - whole-plant (whole-plant trait): cellular changes integrate into plant-level phenotypes. This is the "measurable" level - phenotypes that can be observed, quantified, and evaluated. Example: OsBADH2 knock-out at the molecular level → loss of enzyme function at the cellular level → 2AP accumulation and aroma production at the whole-plant level.

Level four - economic (economic outcome): whole-plant phenotypes are "valued" through economic channels. The four main channels are (1) market price premium, (2) yield increase, (3) post-harvest loss reduction, and (4) market expansion (especially exports).

A concrete illustration of this pipeline is CRISPR aromatic rice via OsBADH2 knock-out: molecular alteration (indel in OsBADH2 exon 7) → loss of BADH2 enzyme function → accumulation of 2-acetyl-1-pyrroline (2AP) → Jasmine-like aroma at the whole-plant level →

12-18% export price premium (~US$600/tonne versus ~US$530/tonne) → estimated US$30-50 million additional revenue annually per newly disseminated aromatic variety.

This pipeline analysis has two important implications. First, the economic value of CRISPR/Cas9 does not lie in DNA cleavage - this is merely the initial event. The true value lies in the capacity to consistently translate from the molecular to the phenotypic and from the phenotypic to the economic levels. Translation efficiency depends on many factors beyond CRISPR itself: biological understanding (choosing the right target genes), plant transformation, tissue culture and propagation systems, variety registration, marketing, and distribution. A country aiming to succeed with CRISPR must invest synchronously across the entire pipeline, not in CRISPR alone.

Second, the three application pathways have distinct "pipeline profiles." NHEJ has a short pipeline (molecular changes rapidly translate into clear phenotypes) - suited to a "quick wins" strategy. HDR has a longer pipeline (donor design, allele verification) but a higher economic value per successful event. dCas9 has a particularly distinctive pipeline - the effect can be "tunable" over time, opening previously unattainable economic models.

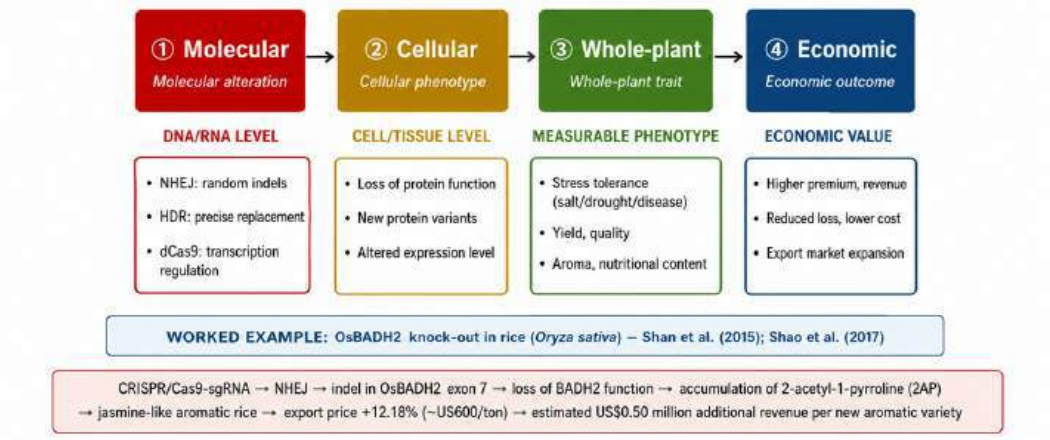


Figure 4. Translation pipeline from molecular alteration to measurable economic value of CRISPR/Cas9 applications in crops. Source: Constructed by the author.

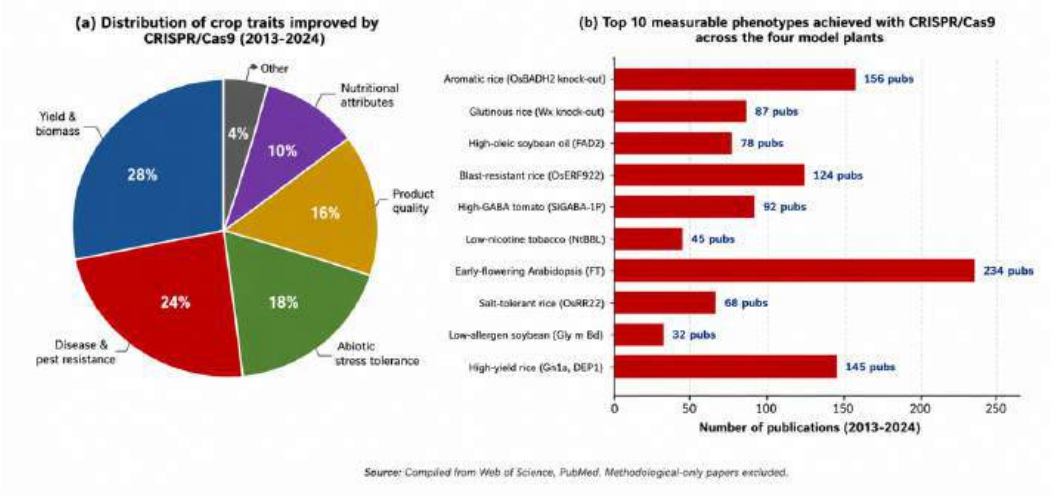


Figure 5. Spectrum of measurable phenotypes produced by CRISPR/Cas9 in model plants (2013-2024). (a) Trait distribution; (b) Top 10 specific phenotypes. Source: Compiled from Web of Science and PubMed.

Table 6. Molecular → cellular → whole-plant → economic translation pipeline - nine illustrative cases

Product	① Molecular	② Cellular	③ Whole-plant	④ Economic	Pathway
Aromatic rice	Indel in OsBADH2	Loss of BADH2	2AP+ → Jasmine aroma	Export premium +12-18%	NHEJ
Blast-resistant rice	Indel in OsERF922	Loss of resistance repressor	Magnaporthe resistance	US$1.5-2.5B/year	NHEJ
Salt-tolerant rice	Indel in OsRR22	Loss of stress repressor	Survival at 150 mM NaCl	Recovers 700,000 ha	NHEJ
High-oleic soybean	Indel in FAD2-1A/B	Loss of conversion	Oleic >80%	Niche US$50-80M	NHEJ
High-GABA tomato	Indel in SlGAD3	Loss of GAD repression	GABA +4-5 fold	Price premium +30-40%	NHEJ
Glyphosate-resistant rice	EPSPS codon replacement	EPSPS resistant to glyphosate	Survives 1500 ppm glyphosate	Potential US$12B market	HDR
Vaccine-expressing tobacco	Cassette knock-in	Stable antigen expression	Plant biopharmaceutical platform	Plant pharming US$4.4B	HDR
Tunable salt-tolerant rice	dCas9 + OsRAV2	OsRAV2 +15-22 fold	Tunable salinity tolerance	Novel smart-agriculture model	dCas9
Arabidopsis apomixis (POC)	dCas9 + AtFIS2	AtFIS2 +25-30 fold	Apomixis proof-of-concept	Potential revolution in food crops	dCas9

Source: Compiled by the author from the studies cited in Tables 3-5 and additional economic analysis.

Policy implications for Vietnam

Based on the analysis of the three CRISPR/Cas9 application pathways and the molecular → phenotypic → economic translation pipeline, four groups of strategic recommendations for Vietnam are proposed.

Recommendation 1: Focus on NHEJ knock-out in rice and soybean as the strategic "entry point" during 2025-2030. The rationale: (i) NHEJ has the highest technical efficiency, matching the capability of existing laboratories (Borphukan et al.); (ii) rice and soybean are the two most strategically important crops for Vietnam; (iii) SDN-1 products derived from NHEJ enjoy regulatory advantages in the main export markets (United States, Japan, United Kingdom, the Philippines) because they do not require a repair template and are often indistinguishable from naturally occurring mutations (Ahmad et al.). Specific flagship projects proposed: (a) Vietnamese CRISPR aromatic rice via OsBADH2 knock-out, using local aromatic rice varieties (ST24, ST25) as backgrounds; budget: US$5-7 million over 4-5 years. (b) Salinity-tolerant rice for the Mekong Delta via knock-out of negative regulators OsRR22 or OsHKT2;3 - budget US$7-10 million. (c) Low-allergen soybean for the Japan-Korea market - budget US$4-6 million.

Recommendation 2: Invest selectively in HDR through international collaboration rather than by building independent capacity. The rationale is that HDR demands very large R&D investment (low efficiency, advanced equipment) and very high technical capability because the delivery of a donor repair template simultaneously with the Cas enzyme remains a severe rate-limiting bottleneck in plant cells (Borphukan et al.). Vietnam should partner with one or two leading global HDR centers (Cold Spring Harbor, KAIST, RIKEN) for technology transfer and training, rather than building HDR capacity from scratch. Priority HDR targets include replacement of the OsEPSPS allele to produce glyphosate-resistant rice (if consistent with the national herbicide strategy) (Borphukan et al.) and insertion of vaccine antigen expression cassettes into tobacco as a biopharmaceutical production platform.

Recommendation 3: Invest early and substantially in dCas9 - first-mover advantages are still available. Currently, only 5% of global plant CRISPR studies use dCas9; competition in this space is still low. Vietnam should establish a dedicated dCas9 R&D program utilizing advanced transcriptional activation and repression architectures to systematically expand plant traits (Lv et al.). This program should target two objectives: (a) dCas9-VP64 to activate stress-response genes in rice, maize, and vegetables (Lv et al.); (b) dCas9-KRAB to repress genes linked to quality and allergen issues in export crops. The proposed budget is US$6-8 million over 5-7 years. Furthermore, dCas9 enjoys a special regulatory advantage - it does not generate "mutations" in the traditional sense - making it a strategic long-term asset.

Recommendation 4: Develop the regulatory framework and research infrastructure in parallel. (a) Issue a dedicated decree for plant genome-editing technology within 18-24 months, distinguishing clearly between SDN-1 (NHEJ knock-out), SDN-2 (small HDR replacements), SDN-3 (large HDR insertions), and dCas9 (Ahmad et al.). (b) Invest in internationally compliant BSL-2 laboratories at 5-7 hubs (VAAS, AGI, VNU Hanoi, VNU HCMC, Can Tho University,

Hue University). (c) Establish a "CRISPR Specialist Pipeline" scholarship program sending 30-50 PhD and postdoctoral researchers abroad annually with a focus on plant CRISPR. The total budget for the regulatory framework and infrastructure is estimated at US$25-30 million over 5-7 years.

Total budget for all four recommendation groups: US$50-70 million over 5-7 years - equivalent to 0.03-0.04% of central government annual expenditures during the same period. This is a feasible investment level when weighed against the potential economic benefits. The Vietnamese CRISPR aromatic rice alone, if it captures 20% of the 2030 aromatic-rice export market (estimated at US$2.5-3 billion), could generate US$80-120 million in additional value annually.

Finally, a key structural factor is the "path-to-market." Vietnam must synchronously build all four levels of the molecular → phenotypic → economic pipeline: (1) molecular - CRISPR-standard laboratories; (2) cellular - efficient tissue-culture and regeneration systems; (3) whole-plant - field-trial networks with strong genetic backgrounds; (4) economic - variety registration, marketing, and export-market integration systems. Heavy investment in CRISPR alone, neglecting other levels, will fail to generate economic value.

Table 7. Proposed ten flagship CRISPR/Cas9 projects for Vietnam (distribution across the three pathways)

Flagship	Pathway	Crop	Gene	Budget	Years
FL1: CRISPR-VN aromatic rice	NHEJ	Rice (ST24/25)	OsBADH2	US$5.5M	4
FL2: Mekong salt-tolerant rice	NHEJ	OM rice varieties	OsRR22/HKT2; 1	US$8.5M	5
FL3: Low-allergen soybean	NHEJ	Soybean (DT)	Gly m Bd	US$5.0M	4
FL4: Blast-resistant rice	NHEJ	Rice	OsERF922	US$4.0M	3
FL5: Glyphosate-tolerant rice (int'l)	HDR	Rice	OsEPSPS	US$6.0M	6
FL6: Vaccine-expressing tobacco	HDR	Tobacco	Safe harbor	US$4.5M	5
FL7: Tunable salt-tolerant rice	dCas9 (CRISPRa)	Rice	OsRAV2	US$5.0M	5
FL8: Soybean oil quality (dCas9)	dCas9 (CRISPRi)	Soybean	GmFAD2 paralogs	US$4.5M	5
FL9: BSL-2 laboratory infrastructure	All	-	-	US$12.0M	5
FL10: CRISPR human-resource training	All	-	-	US$8.0M	7
TOTAL	-	-	-	US$63.0M	5-7

Source: Compiled from established gene-editing targets in agricultural literature including Shan et al. 791; Sheng et al. 8025; and Wang et al. e0154027. Budget distribution by pathway: NHEJ knock-out ~36.5%; HDR replacement ~16.7%; dCas9 transcriptional ~15.1%; infrastructure + training ~31.7%.

Conclusion

This paper has presented a systematic review of the three functional application pathways of CRISPR/Cas9 in plant biology - gene knock-out via NHEJ, targeted sequence replacement via HDR, and transcriptional control via dCas9 - examined across four model species: Arabidopsis

thaliana, rice (Oryza sativa), tobacco (Nicotiana tabacum), and soybean (Glycine max). Drawing on studies from 2013 to 2024, the paper has systematised the translation from molecular alteration to measurable phenotype and quantified the economic value of the achieved phenotypes.

Key findings may be summarised as follows. First, NHEJ is the most common pathway (>85% of studies), with high efficiency (50-90%), favourable regulatory treatment, and many commercial products - making it an ideal "entry point" for countries new to CRISPR applications. Second, HDR offers high economic value per event but suffers from very low plant efficiency (0.1-5%) - a technical bottleneck that must be resolved through long-term R&D or international collaboration. Third, dCas9 opens a new paradigm of expression regulation without DNA modification, with substantial long-term economic potential and first-mover advantages still available.

Fourth, the molecular → cellular → whole-plant → economic translation pipeline is the essential analytical framework for understanding the economic value of CRISPR/Cas9. The true value does not reside in the initial DNA-cleavage event but in the consistent translation across these four levels. Effective R&D investment must be synchronised across the entire pipeline, not concentrated solely on CRISPR. Fifth, for Vietnam, the optimal strategy is to begin with NHEJ in rice and soybean, pursue HDR through international collaboration, and invest early in dCas9 - total budget US$50-70 million over 5-7 years, with the potential to add hundreds of millions of US dollars per year to the agricultural sector.

The paper has several limitations. As a theoretical review, it is not an empirical study based on primary data; economic estimates are drawn from the literature and market reports. The paper focuses on classical CRISPR/Cas9; variants such as Cas12a, Cas13, base editors, and prime editors have distinct application contexts and warrant separate treatment. Future research should focus on (i) empirical investigation of the efficiency of the three pathways across indigenous Vietnamese rice and soybean varieties; (ii) econometric analysis of the impact of flagship projects upon implementation; and (iii) consumer acceptance and willingness-to-pay studies for CRISPR products.

In the context of Vietnam's transition from a resource-based to a knowledge-based agriculture, the three CRISPR/Cas9 application pathways - NHEJ, HDR, dCas9 - are not isolated technical tools, but three strategic fronts with distinct economic profiles. Understanding these differences is the prerequisite for designing scientifically rigorous, reasonable, and strategically long-term public policy in agricultural biotechnology.

Works Cited

Ahmad, A., et al. “An Outlook on Global Regulatory Landscape for Genome-Edited Crops.” International Journal of Molecular Sciences, vol. 22, no. 21, 29 Oct. 2021, p. 11753, 10.3390/ijms222111753.

Belhaj, K., et al. "Editing Plant Genomes with CRISPR/Cas9." Current Opinion in Biotechnology, vol. 32, 2015, pp. 76-84.

Borphukan, B., et al. “A Gemini Virus-Derived Autonomously Replicating System for HDR-Mediated Genome Editing of the EPSP Synthase Gene in Indica Rice.” Mdpi.Com, 2026, https://www.mdpi.com/2223-7747/14/3/477.

Bortesi, L., and R. Fischer. "The CRISPR/Cas9 System for Plant Genome Editing and Beyond." Biotechnology Advances, vol. 33, no. 1, 2014, pp. 41-52.

Cai, Y., et al. "CRISPR/Cas9-mediated Targeted Mutagenesis of GmFT2a Delays Flowering Time in Soybean." Plant Biotechnology Journal, vol. 16, no. 1, 2018, pp. 176-85.

Cai, Y., et al. "Mutagenesis of GmFT2a and GmFT5a Mediated by CRISPR/Cas9 Contributes for Expanding the Regional Adaptability of Soybean." Plant Biotechnology Journal, vol. 18, no. 1, 2020, pp. 298-309.

Chen, M., and Lei S. Q. “Repurposing CRISPR System for Transcriptional Activation.” Advances in Experimental Medicine and Biology, vol. 983, 2017, pp. 147–157, www.ncbi.nlm.nih.gov/pmc/articles/PMC10211423/, https://doi.org/10.1007/978-981-10-4310-9_10.

CJEU. Judgment in Case C-528/16, Confédération Paysanne and Others v Premier Ministre. 2018.

Demorest, Z. L., et al. "Direct Stacking of Sequence-specific Nuclease-induced Mutations to Produce High Oleic and Low Linolenic Soybean Oil." BMC Plant Biology, vol. 16, 2016, p. 225.

Doudna, J. A., and E. Charpentier. "The New Frontier of Genome Engineering with CRISPR-Cas9." Science, vol. 346, no. 6213, 2014, p. 1258096.

Endo, M., et al. "Multigene Knockout Utilizing Off-target Mutations of the CRISPR/Cas9 System in Rice." Plant and Cell Physiology, vol. 56, no. 1, 2015, pp. 41-47.

Endo, M., et al. "Biallelic Gene Targeting in Rice." Plant Physiology, vol. 170, no. 2, 2016, pp. 667-77.

European Commission. Proposal for a Regulation on Plants Obtained by Certain New Genomic Techniques. EC, 2023.

FAO. The State of Food and Agriculture 2022. FAO, 2022.

Fauser, F., et al. "Both CRISPR/Cas-based Nucleases and Nickases Can Be Used Efficiently for Genome Engineering in Arabidopsis Thaliana." The Plant Journal, vol. 79, no. 2, 2014, pp. 348-59.

Gao, J., et al. "CRISPR/Cas9-mediated Targeted Mutagenesis in Nicotiana Tabacum." Plant Molecular Biology, vol. 87, no. 1-2, 2015, pp. 99-110.

General Statistics Office of Vietnam. Statistical Yearbook of Vietnam 2023. Statistical Publishing House, 2024.

Gostimskaya, I. "CRISPR-Cas9: A History of Its Discovery and Ethical Considerations." Biochemistry (Moscow), vol. 87, no. 8, 2022, pp. 777-88.

Grand View Research. Agricultural Biotechnology Market Size, Share & Trends Analysis Report 2024-2030. 2024.

Haun, W., et al. "Improved Soybean Oil Quality by Targeted Mutagenesis of the Fatty Acid Desaturase 2 Gene Family." Plant Biotechnology Journal, vol. 12, no. 7, 2014, pp. 934-40.

Jiang, W., et al. "Demonstration of CRISPR/Cas9/sgRNA-mediated Targeted Gene Modification in Arabidopsis, Tobacco, Sorghum and Rice." Nucleic Acids Research, vol. 41, no. 20, 2013, p. e188.

Jinek, M., et al. "A Programmable Dual-RNA-guided DNA Endonuclease in Adaptive Bacterial Immunity." Science, vol. 337, no. 6096, 2012, pp. 816-21.

Li, J.-F., et al. "Multiplex and Homologous Recombination-mediated Genome Editing in Arabidopsis and Nicotiana Benthamiana Using Guide RNA and Cas9." Nature Biotechnology, vol. 31, no. 8, 2013, pp. 688-91.

Li, M., et al. "Reassessment of the Four Yield-related Genes Gn1a, DEP1, GS3, and IPA1 in Rice Using a CRISPR/Cas9 System." Frontiers in Plant Science, vol. 7, 2016, p. 377.

Li, Z., et al. "Cas9-guide RNA Directed Genome Editing in Soybean." Plant Physiology, vol. 169, no. 2, 2015, pp. 960-70.

Liu, X., et al. "Application of CRISPR/Cas9 in Plant Biology." Acta Pharmaceutica Sinica B, vol. 7, no. 3, 2017, pp. 292-302.

Lowder, L. G., et al. "A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation." Plant Physiology, vol. 169, no. 2, 2015, pp. 971-85.

Lowder, L. G., et al. "Robust Transcriptional Activation in Plants Using Multiplexed CRISPR-Act2.0 and mTALE-Act Systems." Molecular Plant, vol. 11, no. 2, 2018, pp. 245-56.

Lv, J., et al. “Research Progress and Applications of Gene Activation Editing Technology in Crops.” Frontiers in Plant Science, vol. 17, 20 Mar. 2026, 10.3389/fpls.2026.1787461.

Ma, X., et al. "A Robust CRISPR/Cas9 System for Convenient, High-efficiency Multiplex Genome Editing in Monocot and Dicot Plants." Molecular Plant, vol. 8, no. 8, 2015, pp. 1274-84.

Mehrotra, S., et al. "Plant Gene Activation Through dCas9 Chromatin Engineering." Plant Methods, vol. 13, 2017, p. 93.

Mercx, S., et al. "Inactivation of the β(1,2)-xylosyltransferase and the α(1,3)-fucosyltransferase Genes in Nicotiana Tabacum BY-2 Cells by a Multiplex CRISPR/Cas9 Strategy." Plant Cell Reports, vol. 36, no. 5, 2017, pp. 745-57.

Ministry of Agriculture and Rural Development of Vietnam. Statistical Yearbook of Vietnamese Agriculture 2023. 2024.

Ministry of Science and Technology of Vietnam. Annual Report on Agricultural Science and Technology 2020-2024. 2024.
Mordor Intelligence. Agricultural Genome Editing Market Outlook 2025-2030. 2025.
Nelson, Richard R., and Sidney G. Winter. An Evolutionary Theory of Economic Change. Belknap Press, 1982.
OECD. Science, Technology and Innovation Outlook 2023. OECD Publishing, 2023.
Page, Matthew J., et al. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." BMJ, vol. 372, 2021, p. n71.
Park, J. J., et al. "CRISPR/dCas9-mediated Activation of Multiple Endogenous Genes in Arabidopsis." Plant Methods, vol. 13, 2017, p. 89.
Phillips McDougall. Herbicide Tolerant Crops Market Report. 2024.
Piatek, A., et al. "RNA-guided Transcriptional Regulation in Plants Via Synthetic dCas9-based Transcription Factors." Plant Biotechnology Journal, vol. 13, no. 4, 2015, pp. 578-89.
Porter, Michael E. Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, 1985.
Quach, N. A., and H. H. Le. "Current Status and Orientations for the Development of Agricultural Biotechnology in Vietnam." Vietnam Journal of Agriculture and Rural Development, vol. 22, no. 3, 2022, pp. 5-12.
Richardson, Christopher D, et al. “Enhancing Homology-Directed Genome Editing by Catalytically Active and Inactive CRISPR-Cas9 Using Asymmetric Donor DNA.” Nature Biotechnology, vol. 34, no. 3, 20 Jan. 2016, pp. 339–344, https://www.nature.com/articles/nbt.3481, 10.1038/nbt.3481.
Schachtsiek, J., and F. Stehle. "Nicotine-free, Nontransgenic Tobacco (Nicotiana Tabacum L.) Edited by CRISPR-Cas9." Plant Biotechnology Journal, vol. 17, no. 12, 2019, pp. 2228-30.
Schiml, S., et al. "The CRISPR/Cas System Can Be Used as Nuclease for In Planta Gene Targeting and as Paired Nickases for Directed Mutagenesis in Arabidopsis." The Plant Journal, vol. 80, no. 6, 2014, pp. 1139-150.
Schmidt, S. M., et al. "The Evolving Landscape Around Genome Editing in Agriculture." EMBO Reports, vol. 21, no. 6, 2020, p. e50680.
Schumpeter, Joseph A. Capitalism, Socialism and Democracy. Harper, 1942.
Shan, Q., et al. "Genome Editing in Rice and Wheat Using the CRISPR/Cas System." Nature Protocols, vol. 9, no. 10, 2015, pp. 2395-410.
Shan, Q., et al. “Creation of Fragrant Rice by Targeted Knockout of theOsBADH2gene Using TALEN Technology.” Plant Biotechnology Journal, vol. 13, no. 6, 20 Jan. 2015, pp. 791–800, 10.1111/pbi.12312.
Shao, G., et al. "CRISPR/CAS9-mediated Editing of the Fragrant Gene Badh2 in Rice." Chinese Journal of Rice Science, vol. 31, no. 2, 2017, pp. 216-22.
Sheng, X., et al. “Novel Salinity-Tolerant Third-Generation Hybrid Rice Developed via CRISPR/Cas9-Mediated Gene Editing.” International Journal of Molecular Sciences, vol.

24, no. 9, 1 Jan. 2023, p. 8025, https://www.mdpi.com/1422-0067/24/9/8025, 10.3390/ijms24098025.
Shimatani, Z., et al. "Targeted Base Editing in Rice and Tomato Using a CRISPR-Cas9 Cytidine Deaminase Fusion." Nature Biotechnology, vol. 35, no. 5, 2017, pp. 441-43.
Smyth, Stuart J. "Regulatory Barriers to Improving Global Food Security." Global Food Security, vol. 26, 2020, p. 100440.
Sugano, S. S., et al. "Soybean Genome Editing Using CRISPR-Cas9 Technology and Its Application." Breeding Science, vol. 70, no. 4, 2020, pp. 408-20.
Sun, Y., et al. "Engineering Herbicide-resistant Rice Plants Through CRISPR/Cas9-mediated Homologous Recombination of Acetolactate Synthase." Molecular Plant, vol. 9, no. 4, 2016, pp. 628-31.
UN. World Population Prospects 2022. United Nations, 2022.
USDA. Movement of Certain Genetically Engineered Organisms - SECURE Rule (7 CFR Part 340). USDA APHIS, 2020.
Wang, F., et al. "Enhanced Rice Blast Resistance by CRISPR/Cas9-targeted Mutagenesis of the ERF Transcription Factor Gene OsERF922." PLoS One, vol. 11, no. 4, 2016, p. E0154027.
World Bank. Vietnam Agricultural Sector Brief: Climate Smart Agriculture. 2023.
Zhang, A., et al. "Enhanced Rice Salinity Tolerance Via CRISPR/Cas9-targeted Mutagenesis of the OsRR22 Gene." Molecular Breeding, vol. 39, no. 3, 2019, p. 47.
Zhang, Y., et al. "Applications and Potential of Genome Editing in Crop Improvement." Genome Biology, vol. 19, 2018, p. 210.

Who Benefits from AI? Operational Efficiency and Cost Pass-Through in Quick Commerce Platforms By Tanishkka Aggarwal

1. Introduction

Quick commerce platforms are online services that deliver groceries and daily essentials in under 30 minutes. They have grown rapidly in recent years due to networks of hyperlocal dark stores and logistics systems that would have seemed implausible a decade ago. Valued at approximately $199.92 billion globally in 2026, this market is projected to reach $385.36 billion by 2034 with a CAGR of 8.55%. This growth is driven by the use of artificial intelligence (AI), which is important to the platforms' operational capabilities and creates substantial efficiency gains. In fact, AI-enabled supply chain operations achieve around 15% lower logistics expenses and 35% lower inventory levels (McKinsey Global Institute, as cited in Mohsen, 2023).

While AI undoubtedly enhances operational performance, one must ask who ultimately benefits from these technological gains? Is it consumers through lower prices, or platforms through higher margins? This question is particularly important in quick commerce, as it typically operates in oligopolistic markets where firms can afford to retain cost savings as margin. As a result, reductions in marginal costs do not automatically translate into lower prices, raising concerns about the concentration of gains from technological progress.

The economic literature on cost pass-through shows that market structure plays a major role. In highly competitive markets, firms that hold onto cost savings risk losing customers to rivals who do not. This incentivises them to pass on savings, leading to a pass-through rate of around 1.0. Firms in concentrated markets like oligopolies or monopolies, however, can keep prices constant even as costs fall, capturing efficiency gains as profit. So, pass-through rates in such industries are estimated to be between 0.3 and 0.7. Quick commerce lies in the latter category, as there is market power coming from data advantages, algorithmic optimisation, and scale economies, all of which can weaken competitive pressure. Thus, AI may strengthen platform efficiency without necessarily improving consumer welfare through lower prices.

This paper investigates whether AI-enabled operational efficiency gains in quick commerce are passed through to consumers or captured by platforms as higher margins. Specifically, the degree of cost pass-through is being measured, which is the extent to which reductions in costs faced by the firm are reflected in consumer pricing. By using publicly available financial statements from DoorDash, Instacart, and the Indian platforms Blinkit (under Zomato Eternal) and Swiggy Instamart, the analysis constructs baseline and incremental markup measures from per-order revenue and cost data to assess whether cost reductions are reflected in proportional (full), less-than-proportional (partial), or negligible (zero) revenue changes. Variations across market structures and stages of platform development will also be considered.

This research has several implications, even beyond the industry. AI is becoming increasingly embedded in the operational infrastructure of digital markets, and the question of who benefits from its efficiency gains is not unique to quick commerce. If platforms capture these gains instead of passing them through to consumers, it challenges the assumption that

technological progress inherently serves the interests of everyone, and it raises questions about what regulatory frameworks are needed around platform markets. This paper offers empirical grounding for a debate that will only matter more as AI spreads across digital markets.

2. Conceptual Framework

2.1 Economic Theory of Cost Pass-Through

The relationship between cost reductions and consumer prices is not automatic but mediated by market structure and competitive intensity. In perfectly competitive markets, firms are price-takers with no influence over market prices. When production costs fall, competitive pressure forces firms to cut prices proportionally to avoid losing market share to rivals who pass savings through. The pass-through rate is thus close to unity, meaning a one-dollar reduction in marginal cost translates into a one-dollar price reduction since a firm trying to retain cost savings as profit would be undercut by competitors, causing a loss of customers and revenue.

Oligopolies work differently. When a small number of firms dominate a market, there is interdependence, as each price must reflect a firm's own cost structure and how rivals would respond. This disincentives aggressive price competition. Firms recognise that cutting prices may trigger retaliatory cuts, reducing industry-wide profitability without giving any sustainable competitive advantage (Wang & Werning, 2022). The rational response is often to adopt sticky prices and capture any cost savings as increased profit margins. Empirical estimates of pass-through rates in such concentrated markets typically fall between 0.3 and 0.7, well below unity expected under competition (Weyl & Fabinger, 2013; Zimmerman & Carlson, 2010).

Market power reinforces this. A firm with market power faces a downward-sloping demand curve, meaning it can price above marginal cost without losing all its customers. When such a firm cuts costs, the profit-maximising response depends on price elasticity of demand. If consumers are relatively unresponsive to price changes because switching costs are high or alternatives are limited, the firm gains more by retaining the savings than by cutting prices to expand volume (Hoch et al., 1995). So, cost reductions and price reductions decouple and the efficiency gains are retained by the firms not passed through to consumers. This decoupling is not a market failure, but a rational response to the incentives market power creates.

2.2 Platform Economics and Market Power

Quick commerce platforms have several characteristics that concentrate market power and reduce competition. Firstly, network effects create a self-reinforcing cycle. As more consumers use a platform, it becomes more attractive to delivery riders and merchants, and as more riders and merchants join, the platform becomes more valuable to consumers through wider selection and faster delivery (Rochet & Tirole, 2003). These indirect network effects create barriers to entry and allow incumbent platforms to maintain dominant positions even when new entrants offer superior services (Katz & Shapiro, 1985).

Secondly, economies of scale and density add to this. Fixed costs, including technology infrastructure, brand marketing, and dark store establishment, are spread across a larger transaction base as platforms grow. Also, high order volume in a given neighbourhood, also known as operational density, allows a platform to operate dark stores closer to customers. New entrants cannot replicate these localised network effects without achieving critical mass, creating a competitive moat around incumbent platforms (Evans & Schmalensee, 2016).

Thirdly, switching costs, while seemingly trivial in digital markets, also play an important role. Consumers build up transaction histories, learn interfaces, and get loyalty rewards, all of which raise the effective cost of moving to a rival platform (Farrell & Klemperer, 2007; Klemperer, 1995). These factors reduce the price sensitivity of existing customers, allowing platforms to keep similar prices even when cost structures improve.

These factors together create an oligopolistic market structure which characterises most quick commerce markets. The platforms possess enough market power to sustain prices despite operational improvements, and the competitive dynamics do not force aggressive pass-through of cost savings (Jullien & Sand-Zantman, 2021).

2.3 AI-Driven Operational Efficiency

The efficiency gains that AI generates in quick commerce work through many channels, which address inefficiencies in traditional logistics and inventory management.

Firstly, demand forecasting has become more accurate. Machine learning models analyse historical transactions, seasonality, weather, local events, and real-time browsing behaviour to predict product demand at a hyperlocal level. This aligns stocks with demand, reducing stockouts that cause lost sales and overstocking that causes spoilage and markdowns, ultimately leading to a lower cost of goods sold per order. Secondly, route optimization algorithms process real-time traffic, delivery locations, and rider availability data to determine the most efficient delivery sequences, while AI-enabled order batching groups spatially compatible deliveries into single trips. This reduces average delivery times, lowers fuel and vehicle maintenance costs, and increases the number of orders each rider can complete per shift, thus reducing per-order labour costs. Thirdly, workforce scheduling uses predictive analytics to match rider supply with anticipated demand across times of day, reducing idle time in slow periods without under-staffing during peaks. This lowers labour costs per order without compromising service quality or delivery speed. Finally, inventory placement algorithms stock dark stores by localised demand and dynamically rebalance stock between locations as demand shifts. Positioning high-demand items closer to customers reduces delivery distances, delivery times and transportation costs. This enhances efficiency by reducing excess inventory in low-demand areas and prevents stockouts in high-demand ones.

Together, these AI-enabled optimisations reduce cost per order and improve contribution margins, which is the difference between revenue and variable costs per transaction. Importantly, unlike a reduction in supplier prices that might be advertised, or a change in delivery fee that appears explicitly in checkout, improvements in route efficiency or forecast accuracy are

invisible to customers. This opacity matters because it creates information asymmetry, giving platforms discretion over whether to pass savings through or retain them as margin expansion, without facing public or competitive pressure to do either.

2.4. The Pass-Through Decision

When a platform achieves cost reductions through AI, it faces a strategic choice about how to allocate these gains. Full pass-through to consumers, where prices fall proportionally, maximises consumer welfare and can expand market share if competitors do not match the price cut. But the firm misses out on potential margin expansion and risks triggering a price war that reduces profitability industry wide. Partial pass-through acts as a compromise where some savings are shared with consumers through small price reductions while retaining a portion as improved margins. Zero pass-through, which is maintaining prices despite lower costs, maximises short-term profitability by capturing the entirety of the efficiency gains as margin, at the risk of losing market share if competitors price more aggressively.

The optimal strategy depends on several factors. The degree of market competitiveness determines the degree to which rivals can punish pricing decisions that deviate from the market price. In competitive markets with multiple platforms of around equal market power, holding prices above efficient levels is difficult to sustain due to aggressive price competition. In concentrated markets with a dominant firm, competitive pressure is weaker and the platform can retain cost savings more freely without losing much market share (Amiti et al., 2019).

The growth phase of the market matters as well. Platforms in early, loss-making stages of expansion are already pricing below cost to acquire customers, prioritising market share over profitability. At this stage, cost reductions are more likely to be retained to reduce cash burn than to be passed through to consumers who are already receiving below-cost pricing. As markets mature, the pass-through strategy may change, which is investigated by this paper.

The opacity of AI-driven efficiency gains is also important. Visible cost reductions come with competitive and reputational pressures that encourage some pass-through. However, in the case of opaque operational improvements, like AI efficiency gains, platforms face little of such pressure, making it easier to retain these savings without public or competitive backlash.

Finally, a high level of regulatory and public scrutiny for exploitative pricing may push platforms to pass more savings through to showcase consumer benefit from technological progress. Conversely, limited regulatory oversight allows easier margin retention.

3. Literature Review

This review synthesises three interconnected bodies of research, including empirical evidence on AI-driven productivity improvements in logistics, theoretical and empirical work on cost pass-through in oligopolistic markets, and research on platform pricing behaviour and consumer welfare in digital markets, before identifying the gap this paper aims to fill.

3.1 AI and Operational Efficiency in Logistics

There is now substantial evidence that AI adoption meaningfully improves efficiency and cuts costs in logistics and supply chain operations. McKinsey estimates that AI-enabled supply chain operations achieve 15% lower logistics expenses and 35% lower inventory levels (Mohsen, 2023). Industry analyses point to potential cost reductions of 10-25%, across last-mile delivery, warehouse management, and sorting, with aggregate improvements in earnings before interest and taxes of 1-2% (Oliver Wyman Forum, 2024). Other studies show route optimisation algorithms reducing delivery times by 25-40%, AI-driven warehouse systems improving throughput by 40%, and firms with higher AI integration achieving measurably lower operational costs overall (McKinsey Global Institute, 2023; Honeywell, 2024; IRJMETS, 2025).

What is striking is that, despite these documented efficiency gains, relatively few studies ask where these savings actually go, whether they translate into consumer pricing or profit margins. The focus is consistently on operational metrics such as delivery times, forecast accuracy and inventory turnover, leaving the distributional question largely unaddressed.

3.2 Cost Pass-Through in Oligopolistic Markets

The economic literature provides theoretical foundations for understanding when firms retain cost savings rather than passing them through. Classical economic theory establishes that pass-through rates depend a lot on market structure. In perfectly competitive markets, firms pass cost reductions fully to consumers because retaining savings while rivals cut prices results in lost market share. This creates a pass-through rate that approaches unity (Weyl & Fabinger, 2013). However, in oligopolistic markets, firms with significant market power face a choice when costs fall, and the profit-maximising response is often to keep prices similar. Empirical research consistently confirms incomplete pass-through of cost reductions in these settings. Wang and Werning (2022) shows that firms with a market share of 12.5% have cost pass-through rates of approximately 0.5, meaning that only half of cost reductions are reflected in price cuts. Pass-through rates in oligopolies typically sit between 0.3 and 0.7, well below the full pass-through benchmark of unity (Alexeeva-Talebi, 2010; Zimmerman & Carlson, 2010).

The mechanism behind incomplete pass-through in oligopolies is also strategic. Firms in concentrated markets adjust markups in response to cost changes rather than simply passing costs through because they recognise that aggressive price competition may trigger retaliatory responses from rivals that reduce industry-wide margins without generating sustainable competitive advantage (Amiti et al., 2019). Hoch et al. (1995) also shows that the optimal pass-through strategy depends on the curvature of demand, product differentiation, and the nature of competition, with many realistic oligopolies producing pass-through rates below unity.

One finding from this literature is particularly relevant to the quick commerce context. Firm-specific cost reductions, such as AI-driven operational improvements, where only one firm achieves a cost reduction, tend to produce weaker pass-through than industry-wide cost changes (Hepburn et al., 2005). A single firm's efficiency gains are more easily absorbed as margin, especially when the improvement is not observable to competitors or consumers.

3.3 Platform Pricing and Consumer Welfare in Digital Markets

There are certain structural features that explain why digital platforms have a tendency for concentrated market power. Section 2.2 already lays out why network effects, density economies, and switching costs concentrate market power in quick commerce specifically.

Recent platform economics literature also looks at what that concentration leads to on the pricing side. Calvano et al. (2020) show that algorithmic pricing systems can allow for tacit collusion and reduce price dispersion, potentially reducing competition even without explicit coordination. The opacity of these systems matters, as consumers and regulators cannot easily assess whether observed prices reflect competitive conditions or strategic margin retention.

The behavioral economics literature adds that platforms can exploit consumer biases and information asymmetries to extract value through complex pricing structures, hidden fees, and personalized pricing strategies (Armstrong & Vickers, 2012; Gabaix & Laibson, 2006). Studies on consumer behavior on digital shopping platforms suggest that naïve consumers often make poor choices that search and recommendation systems exploit (Dukes & Liu, 2016).

3.4. Gap in the Literature

Each of these three bodies of literature explains a part of the question this paper asks, but none addresses it directly. The logistics and AI literature documents operational improvements without tracing them through to consumer prices versus margin expansion. The cost pass-through literature develops the theoretical and empirical tools for measuring how cost reductions translate into pricing behaviour, but applies them mainly to traditional industries and manufacturing sectors, with very limited focus on digital platform contexts. The platform economics literature establishes that oligopolistic digital markets deviate from other traditional markets in ways that should suppress pass-through, it does not examine AI-driven efficiency gains specifically. What is missing is a study that brings these threads together by asking whether documented AI efficiency improvements in quick commerce are reflected in consumer prices or absorbed into platform margins, and how this varies across platforms at different stages of market development. That gap is what this paper aims to fill.

4. Hypotheses

The previous sections create three testable hypotheses about how AI-driven efficiency gains are distributed in quick commerce, together framing the empirical analysis that follows.

H1: Market Power Hypothesis - AI-enabled efficiency gains are more likely to increase platform margins than lower consumer prices. This hypothesis reflects the oligopolistic nature of quick commerce markets, with dominant platforms having sufficient market power due to network effects, switching costs and scale economies that reduce competitive pressure. In these conditions, margin retention is the profit-maximising response to cost reductions, and platforms have the market power to sustain it without losing significant market share.

H2: Competition Hypothesis - The degree of cost pass-through varies directly with market competitiveness. In markets with more active platforms competing for market share, cost

reductions are more likely to reach consumer prices, since platforms cannot easily sustain prices above the efficient level without risking rivals undercutting them. In markets dominated by a single platform, competitive pressure is weaker and pass-through rates should be lower.

H3: Growth-Phase Hypothesis - During early growth phases, platforms retain efficiency gains to improve unit economics and reduce cash burn rather than passing them through to consumers who are usually already getting subsidised pricing. Pass-through is more likely to happen in later maturity stages when the focus shifts from customer acquisition to sustainable profitability. This hypothesis recognises that changing strategic priorities of platforms across market development stages affects the allocation of cost savings.

These hypotheses are not mutually exclusive. H1 describes the baseline expectation due to market structure, while H2 and H3 describe the conditions under which that baseline might be changed. The following empirical analysis tests which combination of mechanisms best explains observed pricing and margin behaviour across the platforms examined.

5. Data and Methodology

5.1 Research Design

The analysis includes a before-and-after comparison, seeing how platform pricing and margins change across periods of observable operational improvement due to AI. It is difficult to identify the precise moment when AI was adopted since it was a gradual and continuous process rather than a single moment of transformation. Thus, periods where efficiency improvements become visible in the data are instead identified. Metrics such as cost per order reductions and publicly disclosed technology investments are used as proxies for AI adoption.

The comparison has two aspects to it. The first is temporal, by asking how do individual platforms adjust their pricing and margins as operational efficiency improves over time? The second is cross-sectional, by asking is there a difference in pass-through behaviour between markets with different levels of competitive intensity? Together, these two approaches ensure that the analysis can assess all three hypotheses.

Growth phase periods are identified by negative or declining operating margins, high spending on customer acquisition, aggressive promotions, and geographical expansion. Maturity phase periods are identified by slowing geographical expansion, reduced promotions, positive or improving margins, and explicit information on profitability in investor communications. DoorDash and Instacart both showcase the latter characteristics across the analysis period.

5.2 Data Sources

The empirical analysis mainly uses public financial disclosures from quick commerce platforms and their parent companies. For DoorDash and Instacart, annual and quarterly SEC filings provide detailed breakdowns of revenue, cost of goods sold, fulfillment expenses, and operating margins. The 2025 Instacart annual figures are not yet available as a full 10-K and have been estimated by summing the four quarterly 10-Q filings, along with the Q4 2025

shareholder letter where necessary. For the Indian platforms, Zomato's filings with Indian securities regulators cover Blinkit's financials under the Zomato Eternal segment, and Swiggy's post-IPO disclosures provide comparable data for Swiggy Instamart.

Supplementary evidence on operational performance comes from company announcements, technology blog posts, and industry reports documenting specific implementations of AI and their effects on delivery times, order accuracy and efficiency. This gives qualitative evidence for efficiency improvements identified in the financial data. Also, broader information about the operational trends and competition in the market as a whole has been taken from McKinsey, Oliver Wyman, and specialist logistics analytics providers.

This analysis will cover data from 2019 to 2025, which captures both the rapid expansion phase of quick commerce and more recent shift of focus toward profitability. This range is necessary for assessing the Growth-Phase Hypothesis, as it covers different strategic periods for most of the platforms tested.

5.3 Measuring Cost Pass-Through Through Markup Analysis

The main aim is to examine whether cost reductions translate into revenue changes through markup analysis. The markup, defined as revenue per dollar of cost, serves as a proxy for the effective price level relative to underlying costs. Tracking how markups change as costs change allows us to infer how efficiency gains are allocated between consumers and platforms.

The baseline markup for a given platform in period t is calculated as $Markup = \frac{Revenue}{Unit\ cost}$ where total costs include cost of goods sold, fulfillment expenses, delivery labour, and other variable costs directly connected to order fulfillment.

To understand the allocation of each revenue dollar, total revenue is decomposed as $Revenue = Total\ costs + Profit\ contribution$**.** This shows how each dollar of revenue is distributed between covering expenses and generating profit. When costs decline but revenue remains constant, profit contribution increases proportionally, signalling that efficiency gains have been retained rather than passed through.

Pass-through is assessed through the incremental markup, which measures how revenue responds to cost changes between periods through $Incremental\ markup = \frac{\Delta\ Revenue}{\Delta\ Total\ costs}$ where Δ denotes the change between consecutive periods.

Three outcomes are possible. Full pass-through occurs when revenue changes proportionally with costs, leading to an incremental markup close to unity, indicating that the platform cuts effective prices in line with cost savings and margins remain stable. Partial pass-through occurs when revenue falls by less than costs, with the incremental markup being greater than unity, suggesting that savings are shared between consumers and the firm. No pass-through occurs when revenue is stable despite falling costs, suggesting that the platform captures all efficiency gains as margin expansion.

To separate price effects from volume effects, the analysis looks at revenue per order and cost per order alongside aggregate data. Revenue per order becomes a direct proxy for the

effective price paid by consumers, while cost per order captures the platform's unit economics and is the main proxy for AI-driven efficiency improvement.

5.4 Limitations

Several limitations are worth acknowledging. Firstly, publicly available financial data is aggregated across markets and product categories, limiting precision at the city or market level. Platforms do not disclose pricing or cost data by area, so the cross-sectional analysis relies on aggregate figures that may hide some variation. Secondly, the markup approach assumes revenue changes reflect pricing adjustments rather than changes in product mix, order composition, or other revenue streams, a limitation that controlling for revenue per order partially addresses but cannot eliminate. Finally, the analysis cannot directly observe platform decision-making as it is inferred from financial outcomes rather than actual managerial intent. Many explanations for low pass-through may coexist such as exploitation of market power, strategic investment in service quality or constraints due to competitive positioning. Despite these limitations, the combination of temporal and cross-sectional variation offers meaningful evidence on how AI-driven efficiency gains are distributed in quick commerce.

6. Results

6.1 Overview of Data and Sample

The quantitative analysis covers DoorDash, Inc. (NASDAQ: DASH) and Maplebear Inc., trading as Instacart (NASDAQ: CART) over 2019-2025. Both platforms file annual reports with the United States Securities and Exchange Commission (SEC), which includes audited income statement data on revenue, cost of revenue, and order volumes. DoorDash is the dominant quick-commerce and food delivery platform in the United States, while Instacart works mainly as a grocery delivery and technology company. Together, they are the most comprehensively documented publicly listed quick-commerce platforms in the Western market.

Indian platforms Blinkit (a subsidiary of Eternal Limited, formerly Zomato) and Swiggy Instamart were considered for inclusion but excluded from the quantitative analysis because both are currently in loss-making growth phases. Blinkit only reached adjusted EBITDA breakeven in March 2024, and Swiggy Instamart recorded a contribution margin of -5.6% in the fourth quarter of FY2025. Applying the incremental markup framework to platforms that are still subsidising consumers may confuse cash-burning with pass-through and give misleading results. Therefore, they are discussed qualitatively later as contextual evidence for the Growth-Phase Hypothesis (H3) and the Competition Hypothesis (H2).

All financial data is sourced from SEC filings. The 2025 DoorDash figures include the Deliveroo acquisition in October 2025. While this could affect comparability, DoorDash has disclosed organic growth rates excluding Deliveroo, which are noted here. Instacart's 2025 full-year figures are derived from three published quarterly 10-Q reports and the fourth-quarter shareholder letter, as the annual 10-K was not available at the time of analysis.

6.2 H1: Market Power Hypothesis: Markup Analysis

6.2.1 DoorDash

Metric	2019	2020	2021	2022	2023	2024	2025*
Revenue ($M)	885	2886	4888	6583	8635	10722	13717
Cost of revenue ($M)	523	1368	2338	3588	4589	5542	6738
Gross margin %	40.9%	52.6%	52.2%	45.5%	46.9%	48.3%	50.9%
Total orders (M)	263	816	1390	1736	2161	2583	≈ 3172
Revenue per order ($)	3.36	3.54	3.52	3.79	4.00	4.15	4.32
Cost per order ($)	1.99	1.68	1.68	2.07	2.12	2.14	2.12
Baseline markup (Rev ÷ Cost)	1.69×	2.11×	2.09×	1.84×	1.88×	1.93×	2.04×
Net income / (loss) ($M)	(667)	(461)	(468)	(1368)	(565)	+117	+932

Table 1 - DoorDash financial and operational metrics, 2019-2025. Source: DoorDash 10-K filings, SEC CIK 1792789. All figures are in USD millions unless stated.

Several observations from Table 1 are relevant to the Market Power Hypothesis. Firstly, revenue per order increased steadily from $3.36 in 2019 to $4.32 in 2025, a cumulative rise of 29%. This rise in effective consumer pricing shows that the platform retains efficiency gains rather than sharing them through lower fees or delivery charges. Secondly, cost of revenue per order fell during the pandemic demand surge of 2020-21 as volume reduced per-unit costs, then rose through 2022-23 as Wolt integration costs and labour pressures increased, before declining again to $2.12 in 2025. But throughout this period, revenue per order consistently grew faster than cost per order, widening the per-order margin gap from $1.37 in 2019 to $2.20 in 2025.

The baseline markup, revenue divided by cost of revenue, shows how much revenue the platform retains above its fulfillment cost base. This multiple averaged 1.94× across the analysis period, meaning that for every dollar spent on fulfilment, DoorDash generated roughly $1.94 in revenue. More importantly, the markup has trended upward since 2022, rising from 1.84× to 2.04× by 2025, despite cost per order declining over the same period. If efficiency gains were being passed through to consumers, the markup would be expected to remain stable or fall, as lower costs would be reflected in lower prices. Instead, the rising markup suggests that cost reductions have not translated into lower consumer-facing revenue, and that the gains have been retained as expanded margins, consistent with the Market Power Hypothesis.

Table 2 presents the incremental markup analysis, tracking how each dollar of additional cost is associated with revenue changes across consecutive years. An incremental markup near 1.0× indicates full pass-through, while values well above 1.0× mean revenue grew faster than costs, signalling margin retention.

Period	Δ Revenue ($M)	Δ Cost of revenue ($M)	Incremental markup	Interpretation
2019 → 2020	+2,001	+845	2.37×	Margin retention
2020 → 2021	+2,002	+970	2.06×	Margin retention
2021 → 2022	+1,695	+1,250	1.36×	Partial pass-through
2022 → 2023	+2,052	+1,001	2.05×	Margin retention
2023 → 2024	+2,087	+953	2.19×	Margin retention
2024 → 2025*	+2,995	+1,196	2.50×	Accelerating margin retention

Table 2 - DoorDash incremental markup analysis, 2019-2025. * Includes Deliveroo acquisition effects from October 2025.

The incremental markup exceeded 2.0× in four of the six periods measured. The one exception was the period from 2021 to 2022 (1.36×), when cost growth was driven up by the Wolt acquisition and rising insurance and labour costs. The more recent periods give more important insights. The 2022 to 2023 and 2023 to 2024 incremental markups of 2.05× and 2.19× are also periods of significant AI investment, including DoorDash's "Deep Red" logistics optimisation system and expanded demand forecasting capabilities, both disclosed in its 10-K filings. The 2024 to 2025 figure of 2.50× is the highest in the series and suggests that efficiency gains are being captured as margin at an increasing rate as the platform matures. DoorDash recorded its first full-year net profit of $117 million in 2024, which rose to $932 million in 2025, while revenue per order continued to rise throughout.

These findings are consistent with H1. Cost improvements attributable to AI-enabled operational efficiency are not being reflected in proportional reductions in consumer-facing prices. Instead, the efficiency gains have been largely captured as expanded margins.

Fig 2: DoorDash: revenue per order and cost of revenue per order, 2019-2025 - The shaded area indicates the per-order margin gap. Sources: DoorDash 10-K annual reports, SEC 2020-2026.

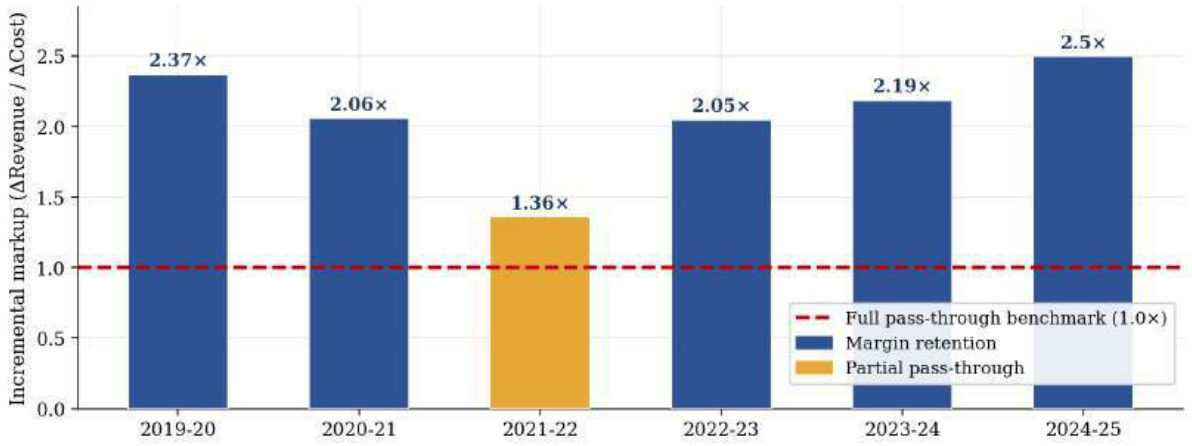


Fig 3: DoorDash incremental markup by period, 2019-2025 - The dashed line indicates the full pass-through benchmark (1.0×). Sources: DoorDash 10-K annual reports, SEC 2020-2026.

6.2.2 Instacart

Metric	2022	2023	2024	2025*
Revenue ($M)	2551	3042	3378	3742
Cost of revenue ($M)	720	764	836	984
Operations & support ($M)	252	344	278	≈ 272
Total operational cost ($M)	972	1108	1114	≈ 1256
Gross margin %	71.8%	74.9%	75.2%	73.7%
Orders (millions)	262.6	269.2	294.0	338.8
Revenue per order ($)	9.71	11.30	11.49	11.04
Cost of revenue per order ($)	2.74	2.84	2.84	2.91
Adjusted EBITDA ($M)	-	641	885	≈ 1087

Table 3 - Instacart financial and operational metrics, 2022-2025. (Pre-2022 data is unavailable due to the platform's private status before IPO in September 2023). Sources: Maplebear Inc. 10-K/10-Q filings (2023-2025). Total operational cost = cost of revenue + operations & support.

Instacart provides the most striking evidence in the analysis. Cost of revenue per order was similar across the four-year period, rising only from $2.74 in 2022 to $2.91 in 2025, representing a 6% increase, while order volumes grew by 29% over the same period. Serving substantially more orders at almost constant per-unit costs clearly shows an increase in operational efficiency. This is consistent with AI-enabled improvements which reduces the marginal cost of an additional order. Instacart's 10-K filings also cite "fulfillment efficiencies" as a driver of financial performance, and the company's Q4 2025 earnings call reported that AI tools had increased average output per engineer by around 40% over the past year.

Despite this stable cost per order, revenue per order grew from $9.71 in 2022 to a peak of $11.49 in 2024, which is an 18% increase. The per-order margin gap therefore widened from $6.97 to $8.65 over this period. Notably, Instacart's revenue per order is significantly higher

compared to DoorDash due to its business model. Around 29% of Instacart's revenue comes from advertising, with brands paying for being prominent in search results and product listings, rather than from delivery fees or service charges paid by consumers. This advertising revenue is invisible to consumers as a pricing mechanism and represents a form of value extraction that is retained by the platform without directly affecting consumers.

In 2025, revenue per order reduced to $11.04. This sudden switch does not actually show consumer pass-through of efficiency gains. As explained in Instacart's shareholder disclosures, the decline was due to strategic decisions to lower basket minimums for Instacart+ members and introduce restaurant orders with lower average values, in order to grow order volumes. This is a volume acquisition strategy, not a price reduction driven by cost improvements, and it is consistent with the company's own framing of the change as an "affordability initiative designed to increase customer engagement."

Period	Δ Revenue ($M)	Δ Total operational cost ($M)	Incremental markup	Interpretation
2022 → 2023	+491	+136	3.61×	Strong margin retention
2023 → 2024	+336	+6	56×	Near-zero cost growth
2024 → 2025*	+364	+142	2.56×	Continued margin retention

Table 4 - Instacart incremental markup analysis, 2022-2025.

The 2022 to 2023 and 2023 to 2024 periods are particularly telling. In both years, revenue growth significantly outpaced cost growth, giving incremental markups of 3.61× and 56× respectively. The extreme 2023 to 2024 figure reflects the fact that total operational costs grew by only $6 million while revenue grew by $336 million, a period that Instacart described in its shareholder communications as showing "operating leverage." This represents a near-textbook case of efficiency gains being retained as margin rather than passed to consumers. The 2024 to 2025 incremental markup of 2.56×, still well above the 1.0× full pass-through benchmark, confirms that this pattern continued into the most recent period.

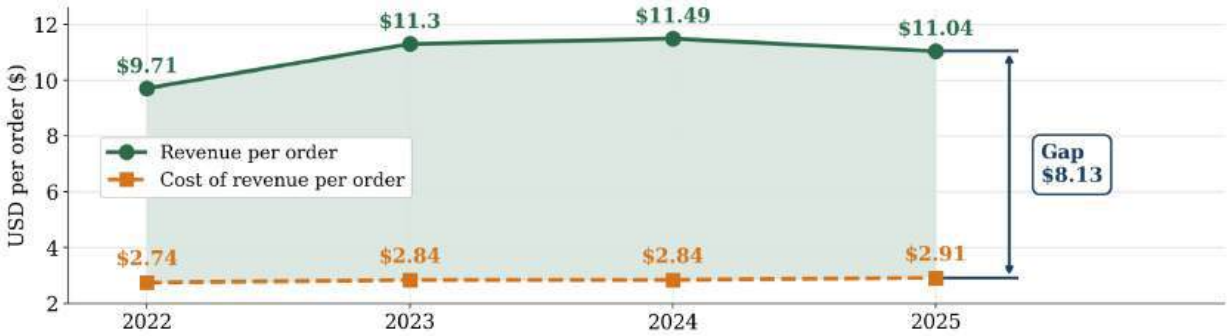


Fig 4: Instacart: revenue per order and cost of revenue per order, 2022-2025 - The shaded area indicates the per-order margin gap. Sources: Maplebear Inc. 10-K/10-Q filings (2023-2025).

6.3 Cross-Platform Comparison and H2: Competition Hypothesis

H2 predicts that pass-through rates should be higher in more competitive markets. The competitive nature of the markets is very different for the two platforms. DoorDash holds approximately 67% market share in US food delivery and operates in a market characterised by two dominant players (DoorDash and Uber Eats). Instacart holds a dominant position in US online grocery, with limited direct quick-commerce competition in most of its markets.

Metric	DoorDash (2024)	Instacart (2024)	Directional implication
Revenue per order ($)	4.15	11.49	Both rising YoY
Cost of revenue per order ($)	2.14	2.84	Both flat/declining
Per-order margin gap ($)	2.01	8.65	Both widening (2022-2024)
Baseline markup	1.93×	4.04×	Both above 1.0× benchmark
Gross margin %	48.3%	75.2%	Both improving or stable
Avg. incremental markup (excl. peak year)	≈ 2.1×	≈ 3.1×	Both well above 1.0×

Table 5 - Cross-platform comparison of per-order efficiency and revenue metrics, 2024 (as it is the most recent full year with complete data for both platforms) - Average incremental markup excludes outlier years (DoorDash 2021-22; Instacart 2023-24).

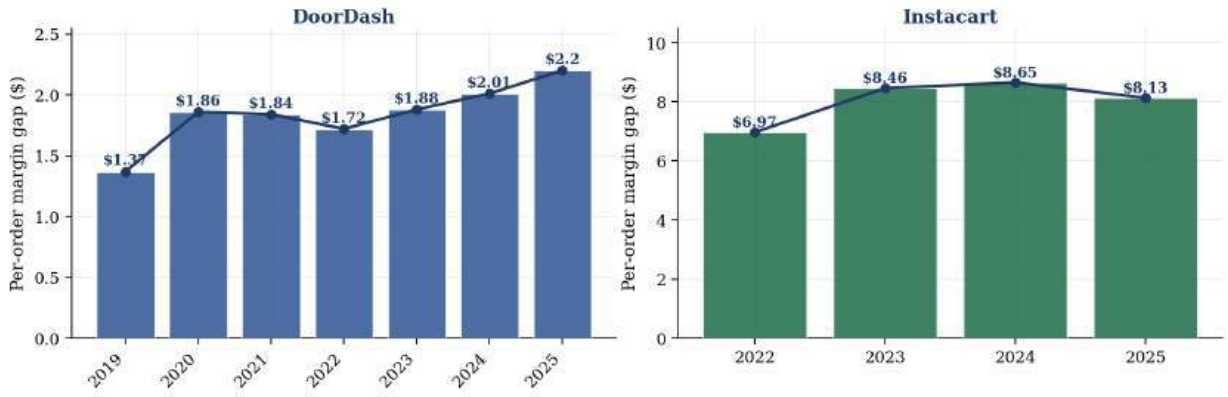


Fig 5: Per-order margin gap, DoorDash and Instacart - Margin gap = revenue per order - cost of revenue per order. A widening gap indicates efficiency gains retained as margin rather than passed through to consumers. Sources: SEC 10-K filings.

Both platforms exhibit baseline markups well above 1.0×, meaning both retain substantial revenue above fulfillment costs, and both show widening per-order margin gaps over the analysis period. This similar behaviour, despite the platforms operating in very different market segments, supports H1. It suggests structural market power matters more in determining pass-through than competitive intensity alone.

A rigorous cross-sectional test of H2 is not possible with this dataset, since there is not enough variation in competitive structure between the two platforms to meaningfully compare them. Instacart's higher margins stem from its business model, not lower competitive intensity.

Some indirect evidence for H2 can be taken from the Indian market. India's quick-commerce sector has 3-4 significant competing platforms, including Blinkit, Swiggy Instamart, Zepto, and Flipkart Minutes, with market share distributed more evenly than in the US. Rather than retaining efficiency gains as margin, Indian platforms have engaged in aggressive consumer discounting, subsidised delivery, and promotional pricing. This is broadly consistent with what H2 predicts in competitive markets. Blinkit's contribution margin only recently reached zero after years of losses, and Swiggy Instamart's contribution margin was still -5.6% in Q4 FY2025. This suggests that competitive pressure is leading to firms adopting below-cost pricing for consumers, which is a form of pass-through, though the mechanism differs from the direct fee reduction that standard pass-through theory expects.

6.4 H3: Growth-Phase Hypothesis: Temporal Analysis

H3 predicts that efficiency gains will be retained during early growth phases to improve unit economics and reduce cash burn, with pass-through potentially emerging only in maturity phases as platforms shift focus toward profitability optimization.

DoorDash's financial history shows growth-phase dynamics. Between 2019 and 2023, the platform operated at a net loss every year, facing losses of over $3.3 billion. During this period, the company's 10-K filings described strategic priorities being market share growth, consumer acquisition, and geographic expansion, all of which are characteristic of the growth phase. Even so, the incremental markup analysis shows that revenue mostly grew much faster than costs during this period. This means any efficiency gains were being directed towards improving unit economics rather than subsidising consumers further beyond the existing losses.

DoorDash’s transition to maturity is clearly marked by its first full-year net profit, $117 million in 2024, followed by $932 million in 2025. Rather than triggering increased pass-through as H3 predicts, this period had the highest incremental markups in the dataset (2.19× in 2023-24 and 2.50× in 2024-25). This suggests that maturity increases margin retention rather than pass-through, inconsistent with H3 but consistent with H1. Platforms with sufficient market power retain efficiency gains as margins regardless of phase of development.

Following a similar trend, Instacart had already achieved operational profitability by 2022, its first year in this dataset. Across the available period, the incremental markup never approached 1.0×, and adjusted EBITDA grew from $641 million in 2023 to approximately $1,087 million in 2025 as AI-driven efficiencies were converted to profit. There is no financial evidence that the platform's maturity encouraged any consumer pass-through.

By contrast, the Indian platforms exemplify the behaviour H3 predicts. Both Blinkit and Swiggy Instamart have noticeably prioritised market share and volume growth over margin, with negative contribution margins showing that consumers are currently receiving below-cost pricing. This is consistent with H3. However, once these platforms achieve the operational scale

and market position that DoorDash and Instacart have, the theoretical prediction that pass-through would increase in the maturity phase does not appear to be supported by evidence from markets that have already reached that stage.

7. Discussion

The results raise three interpretive questions addressed below: Why did market power dominate over competitive and lifecycle dynamics? What does the Instacart advertising model imply for consumer welfare beyond standard pass-through analysis? What does the Indian evidence suggest about the long-run trajectory of competitive markets?

7.1 Market Power as the Dominant Mechanism

The theoretical framework (Section 2) identified four structural characteristics that concentrate market power in quick commerce such as network effects, scale and density economies, switching costs, and algorithmic pricing opacity. The empirical results show all four working together to discourage pass-through.

DoorDash's dominant position in US food delivery, with approximately 67% market share, means that it faces very little competitive threat that would force it to lower prices due to cost improvements. The presence of Uber Eats as a second major platform creates the appearance of competition, but in oligopolistic markets, strategic interdependence works against aggressive price competition rather than for it. Both platforms benefit from the status quo as one firm cutting prices would reduce industry-wide profits without guaranteeing market share gains if the rival matches the cuts. This theory is confirmed by the data as revenue consistently grew faster than costs regardless of what the cost environment was doing.

Instacart's case is structurally similar but the mechanism is slightly different. Its market power does not rest on delivery speed or geographic coverage like DoorDash. It rests instead on its position as the dominant online grocery platform, with relationships with major retailers, an established network of users, and consumer data. These advantages are not easily replicated by a new entrant, and they give Instacart room for setting higher prices. The fact that cost per order barely changed over four years while order volumes grew by nearly 30% demonstrates that AI-enabled efficiency gains were effectively invisible to consumers and rivals, leading to the widening per-order margin gap observed in the data.

Both cases are consistent with the opacity argument in Section 2.4 as consumers cannot observe improvements in fulfilment efficiency, so platforms face no reputational or competitive pressure to reflect those savings in pricing.

7.2 The Instacart Advertising Revenue Problem

Instacart's 29% advertising revenue raises a consumer welfare concern that the markup framework does not fully capture. From a markup perspective, advertising and delivery fee revenue are treated similarly, but from the perspective of consumer welfare, they are quite different. If advertising investment influences which products appear prominently in search

results, it may distort the shopping experience towards higher-margin brands rather than the best-value options. Consumers who believe they are choosing freely from a neutral recommendation system may actually see their choices being shaped by advertiser incentives.

This represents a form of value extraction that does not show up in delivery fee data but is still consequential, as mentioned in literature on platform design and consumer bias (Section 3.3). The Instacart case suggests that this is not just theoretical. As advertising becomes a major revenue stream for quick commerce platforms, the consumer welfare question becomes not just whether delivery fees reflect cost savings, but whether the platform's commercial incentives are aligned with the consumer's interest in finding the best product at the best price.

7.3 Rethinking the Growth-Phase Hypothesis

There are two complementary explanations for why the maturity phase led to faster margin capture rather than greater consumer pass-through. The first is that the transition from growth to maturity does not change the fundamental incentive structure. Market power is a structural feature of the platform's position, not a temporary condition of its growth phase. Once a platform has accumulated the network effects, density advantages, and switching costs that characterise the incumbent, those features persist and continue to support margin retention regardless of whether the platform is still growing aggressively. Profitability does not bring new competitive pressure that did not previously exist.

The second explanation is more specific to the recent nature of quick commerce. DoorDash's transition to profitability occurred in a period of significant AI investment and operational improvement. The platform was reaping efficiency gains that had been accumulating through years of technology deployment. Thus, the maturity phase in this case does not represent a strategic shift toward consumer benefit, but an operational payoff for prior investment. This suggests that the timing of profitability and the timing of highest incremental markups are not coincidental, but causally related.

This has an important implication. H3 may be correct in its description of what happens during the growth phase, that platforms prioritise unit economics over pass-through, but wrong in its prediction about what follows. Both Blinkit and Swiggy Instamart are currently subsidising consumers, maintaining negative or near-zero contribution margins to build market share. H3 correctly identifies this as growth-phase behaviour. But the DoorDash evidence suggests that when these platforms reach maturity, the outcome is more likely to be accelerated margin capture than increased consumer pass-through. So, Indian consumers benefiting from subsidised delivery today should not necessarily expect lower prices in the future. Rather, subsidies will end and prices will normalise around the level the market structure allows.

7.4 Insights About H2 Provided by the Indian Market Evidence

The Indian evidence provides a useful, though indirect, illustration of the mechanisms H2 describes, and it also reveals an important nuance in how pass-through works in competitive markets. Indian platforms are not cutting delivery fees because costs have fallen and competition

forces them to share the savings. They are subsidising consumers because the market has not yet converged, and each platform still needs to buy loyalty. In a structural sense this is pass-through, since consumers are paying below cost, but the mechanism is customer acquisition strategy, not efficient pricing in response to operational improvements.

This complicates the interpretation of H2. The prediction that competitive markets produce higher pass-through is broadly supported, but the mechanism matters. Pass-through in competitive quick commerce markets seems to operate mainly through growth-phase subsidisation rather than through the direct cost-to-price transmission standard theory assumes. The more important question is whether competitive pressure will keep producing consumer benefits once these markets mature, and the DoorDash evidence suggests it may not.

7.5 Implications for Consumer Welfare and Regulation

The evidence suggests that the technological progress in quick commerce is not producing the broad-based consumer benefits that standard economic theory would predict in competitive markets. Rather than receiving lower delivery fees or reduced service charges, consumers are seeing efficiency gains flow to platform shareholders. This challenges the common claim about technology investment, that productivity gains from innovation are distributed through the economy to benefit consumers broadly. In concentrated digital markets, this paper's evidence suggests that the distribution of benefits is limited and slow at best.

The opacity of AI-driven efficiency improvements adds to this problem. In markets where cost reductions are visible and verifiable, consumers and regulators can hold platforms accountable for pricing that does not reflect improved cost structures. But where efficiency gains are buried in operational algorithms, no such accountability exists. As platforms deploy more sophisticated, more opaque AI systems, the gap widens between what platforms know about their own costs and what consumers can infer from outside. As the conceptual framework noted, this information asymmetry is a structural reason for continued margin retention.

For regulators, this raises several questions. The first is whether existing competition frameworks are enough for assessing market power in AI-enabled platform markets. Conventional measures of competitive harm focus on whether prices are above competitive levels. But in markets where costs are falling and prices are stable, the appropriate benchmark is not a static competitive price but a dynamic one that accounts for efficiency gains. A platform holding prices constant while cutting costs is, in effect, raising its real margin even if the nominal price never moves, which may be entirely missed by some regulatory frameworks.

The second question is whether transparency obligations could address the information asymmetry problem. Requiring platforms to disclose per-order cost and revenue data, or to report how AI investments have affected their cost structures, would give consumers, regulators, and researchers the information needed to assess whether efficiency gains are being shared. This may raise concerns about commercial confidentiality and the feasibility of consistent disclosure across platforms. But the alternative, in which AI-driven efficiency gains in concentrated

markets only benefit shareholders while consumers pay fees that do not reflect improving cost structures, raises its own fairness concerns that need serious regulatory attention.

These questions are not unique to quick commerce. As AI becomes embedded in the operations of a growing range of digital markets, the distributional question this paper investigates will come up more often. Quick-commerce offers a clear case study because its unit economics are simple and per-order data is publicly available. In sectors where the economics are less transparent, the problem may be harder to identify but still real.

7.6 Contributions and Limitations

This paper makes two contributions to the existing literature. The first is empirical. By using a per-order markup analysis from public financial disclosures across six years, it provides direct evidence of how AI-driven operational efficiency gains are distributed in quick commerce by being retained as margin, a question the existing literature had not addressed. The second contribution is conceptual. The paper shows that the Growth-Phase Hypothesis may be useful in understanding the behaviour of loss-making platforms, but it does not predict what happens when they mature. The maturity phase does not produce increased pass-through. Instead, market power continues to operate regardless of growth phase. This suggests that the structural features of the market matter more than its stage of development, and that research on platform pricing should prioritise structure analysis over lifecycle analysis.

Nonetheless, a few limitations constrain the conclusions drawn here. The most significant is causal identification. The analysis cannot isolate AI-driven efficiency improvements from other cost changes, and cannot attribute pricing decisions to specific strategic choices rather than other confounding factors such as macroeconomic conditions, input costs, or competitive dynamics not captured in the data. The findings are directional and descriptive rather than causal. They are consistent with the hypothesis that AI efficiency gains are being retained as margin, but one cannot completely rule out alternative explanations.

The dataset is also limited to 2 US platforms, which limits the cross-sectional analysis. A more comprehensive test of H2 would need data across many geographic markets with different competitive structures, which is not publicly available. The Indian platforms do offer qualitative evidence but cannot be analysed quantitatively without undermining the analysis.

Finally, the markup framework captures aggregate platform-level dynamics but cannot distinguish between different types of consumer-facing charges. The data available cannot fully show whether efficiency gains are being retained through stable delivery fees, rising service charges, growing advertising loads, or some combination of these. More granular pricing data, whether from regulatory disclosures, consumer panels, or platform partnerships, would help pin down where exactly the gains are being captured, and at whose expense.

8. Conclusion

This paper asks whether AI-enabled operational efficiency gains in quick commerce are passed through to consumers at lower prices, or retained by platforms with higher margins.

Based on markup analysis of public financial data from DoorDash and Instacart, 2019 to 2025, the answer is clear. Efficiency gains are being retained. Revenue per order rose steadily for both platforms while cost per order stayed flat or declined, and the per-order margin gap widened in every period examined. Incremental markups averaged well above the 1.0× full pass-through benchmark throughout, highest in the most recent and most profitable years. The technological progress within these platforms is real and measurable, but its benefits are accruing primarily to shareholders rather than to the consumers who use the service.

Three hypotheses guided the analysis. H1, the Market Power Hypothesis, predicted that oligopolistic market structure would suppress pass-through and favour margin retention, and the evidence confirms this. H2, the Competition Hypothesis, predicted that more competitive markets would produce higher pass-through, a claim that the Indian market supports only indirectly. Competitive pressure has produced below-cost consumer pricing, though through growth-phase subsidisation rather than direct cost-to-price transmission. H3, the Growth-Phase Hypothesis, predicted that efficiency gains would be retained early and passed through as markets matured. Only the first half holds. DoorDash and Instacart show no pass-through at maturity; instead, maturity brought the dataset's most aggressive margin retention. Market power, not growth phase, is the dominant determinant of pass-through behaviour.

The findings confirm H1, but also raise two further issues. The first issue concerns the advertising revenue pattern at Instacart. As quick-commerce platforms monetise data and consumer attention through advertising, the consumer welfare question is no longer confined to delivery fees and service charges, but also concerns the integrity of the shopping experience itself. Platforms with a commercial incentive to promote certain products may be extracting value from consumers through recommendation systems instead of pricing, a form of margin capture outside conventional pass-through analysis. The second concerns consumers in markets like India, where competition still produces subsidised pricing. Evidence from mature markets suggests consolidation settles not into a competitive equilibrium with efficient pricing, but into an oligopolistic one with sustained margin retention. Consumers benefitting from below-cost pricing today are benefitting from a growth phase that will end, not a market structure built to serve their interests afterward.

These dynamics are not unique to quick commerce. Wherever oligopolistic market power, network effects, and high switching costs accompany substantial AI investment, the outcome is likely to look similar. Firms capture the gains through wider margins first, and consumers see lower prices later, if they see them at all.

Works Cited

Academic and Theoretical Sources

Alexeeva-Talebi, V. (2010). Cost pass-through of the EU emissions allowances: Examining the German electricity market. ZEW Discussion Paper, No. 10-088. Zentrum für Europäische Wirtschaftsforschung (ZEW). https://doi.org/10.2139/ssrn.1735139

Amiti, M., Itskhoki, O., and Konings, J. (2019). International shocks, variable markups, and domestic prices. Review of Economic Studies, 86(6), 2356–2402. https://doi.org/10.1093/restud/rdz005

Armstrong, M. and Vickers, J. (2012). Consumer protection and contingent charges. Journal of Economic Literature, 50(2), 477–493. https://doi.org/10.1257/jel.50.2.477

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. American Economic Review, 110(10), 3267–3297. https://doi.org/10.1257/aer.20190623

Dukes, A. and Liu, L. (2016). Online shopping intermediaries: The strategic design of search environments. Management Science, 62(4), 1064–1077. https://doi.org/10.1287/mnsc.2015.2176

Evans, D. S. and Schmalensee, R. (2016). Matchmakers: The New Economics of Multisided Platforms. Harvard Business Review Press.

Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. In M. Armstrong and R. Porter (Eds.), Handbook of Industrial Organization, Vol. 3 (pp. 1967–2072). Elsevier. https://doi.org/10.1016/S1573-448X(06)03031-7

Gabaix, X. and Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. Quarterly Journal of Economics, 121(2), 505–540. https://doi.org/10.1162/qjec.2006.121.2.505

Hepburn, C., Grubb, M., Neuhoff, K., Matthes, F., and Tse, M. (2005). Auctioning of EU ETS phase II allowances: how and why? Climate Policy, 6(1), 137–160. https://doi.org/10.1080/14693062.2006.9685587

Hoch, S. J., Kim, B.-D., Montgomery, A. L., and Rossi, P. E. (1995). Determinants of store-level price elasticity. Journal of Marketing Research, 32(1), 17–29. https://doi.org/10.1177/002224379503200102

Jullien, B. and Sand-Zantman, W. (2021). The economics of platforms: A theory guide for competition policy. Information Economics and Policy, 54, 100880. https://doi.org/10.1016/j.infoecopol.2020.100880

Katz, M. L. and Shapiro, C. (1985). Network externalities, competition, and compatibility. American Economic Review, 75(3), 424–440. https://www.jstor.org/stable/1814809

Klemperer, P. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. Review of Economic Studies, 62(4), 515–539. https://doi.org/10.2307/2298075

Rochet, J.-C. and Tirole, J. (2003). Platform competition in two-sided markets. Journal of the European Economic Association, 1(4), 990–1029. https://doi.org/10.1162/154247603322493212

Mohsen, B. M. (2023). Impact of artificial intelligence on supply chain management performance. Journal of Service Science and Management, 16, 44–58. https://doi.org/10.4236/jssm.2023.161004

Wang, C. and Werning, I. (2022). Dynamic oligopoly and price stickiness. American Economic Review, 112(8), 2815–2849. https://doi.org/10.1257/aer.20210471

Weyl, E. G. and Fabinger, M. (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. Journal of Political Economy, 121(3), 528–583. https://doi.org/10.1086/670401

Zimmerman, P. R. and Carlson, J. A. (2010). Competition, information, and the sustainability of collusion in the US retail gasoline market. Review of Industrial Organization, 36(4), 321–334. https://doi.org/10.1007/s11151-010-9250-x

Industry and Consultancy Reports

Fortune Business Insights. (2026). Quick commerce market size, share and industry analysis, 2026–2034. Retrieved June 2026 from https://www.fortunebusinessinsights.com/quick-commerce-market-111868

Honeywell. (2024). AI in warehouse operations: Efficiency and productivity gains. Honeywell International Inc. Retrieved from https://www.honeywell.com/us/en/news/2024/ai-warehouse-operations

IRJMETS. (2025). AI adoption in logistics: Delivery time and operational cost reductions. International Research Journal of Modernization in Engineering Technology and Science, 7(1). https://www.irjmets.com

McKinsey Global Institute. (2023). The economic potential of generative AI: The next productivity frontier. McKinsey and Company. Retrieved from https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai

Oliver Wyman Forum. (2024). AI in logistics: Cost reduction and operational transformation. Oliver Wyman. Retrieved from https://www.oliverwyman.com/our-expertise/insights/2024/ai-logistics.html

Precedence Research. (2025). AI in supply chain market size, share and trends, 2025–2035. Retrieved June 2026 from https://www.precedenceresearch.com/ai-in-supply-chain-market

Primary Financial Disclosures

DoorDash, Inc. (2021). Annual report on Form 10-K for the fiscal year ended December 31, 2020. United States Securities and Exchange Commission. CIK 1792789.

https://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=1792789&type=10-K

DoorDash, Inc. (2022). Annual report on Form 10-K for the fiscal year ended December 31, 2021. United States Securities and Exchange Commission. CIK 1792789.

DoorDash, Inc. (2023). Annual report on Form 10-K for the fiscal year ended December 31, 2022. United States Securities and Exchange Commission. CIK 1792789.

DoorDash, Inc. (2024). Annual report on Form 10-K for the fiscal year ended December 31, 2023. United States Securities and Exchange Commission. CIK 1792789.

DoorDash, Inc. (2025). Annual report on Form 10-K for the fiscal year ended December 31, 2024. United States Securities and Exchange Commission. CIK 1792789.

DoorDash, Inc. (2026). Annual report on Form 10-K for the fiscal year ended December 31, 2025. United States Securities and Exchange Commission. CIK 1792789.

All DoorDash SEC filings accessible at: https://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=1792789&type=10-K&dateb=&owner=exclude&count=10

Maplebear Inc. (Instacart). (2024). Annual report on Form 10-K for the fiscal year ended December 31, 2023. United States Securities and Exchange Commission. CIK 1579091.

Maplebear Inc. (Instacart). (2025). Annual report on Form 10-K for the fiscal year ended December 31, 2024. United States Securities and Exchange Commission. CIK 1579091.

Maplebear Inc. (Instacart). (2025). Quarterly reports on Form 10-Q for Q1, Q2, and Q3, fiscal year 2025. United States Securities and Exchange Commission. CIK 1579091.

All Instacart SEC filings accessible at: https://investors.instacart.com/financials-filings/sec-filings

Maplebear Inc. (Instacart). (2026). Q4 2025 shareholder letter. Investor Relations. Retrieved from https://investors.instacart.com

Eternal Limited (formerly Zomato Ltd). (2024). Annual report for the fiscal year ended March 31, 2024. Registrar of Companies, India. Retrieved from https://www.eternal.com/investor-relations/results/

Eternal Limited. (2025). Annual report for the fiscal year ended March 31, 2025. Registrar of Companies, India. Retrieved from https://www.eternal.com/investor-relations/results/

Swiggy Limited. (2025). Annual report for the fiscal year ended March 31, 2024. National Stock Exchange of India. Retrieved from https://www.nseindia.com/get-quote/equity/SWIGGY

Swiggy Limited. (2025). Annual report for the fiscal year ended March 31, 2025. National Stock Exchange of India. Retrieved from https://nsearchives.nseindia.com/annual_reports/AR_27154_SWIGGY_2024_2025_A_29072025154606.pdf

How Accurate is Your Diagnosis? The Effects of COVID-19 on Mental Illnesses & Diagnoses By Yunn Nway Lin

Abstract

In 1977, George Endel, in an effort to improve upon the biomedical model, incorporated psychological and social factors to propose the biopsychosocial model. He did this with the aim of providing a more comprehensive explanation to the various factors that influence health and wellbeing- both the physical and mental aspects of it. While there have been studies conducted venturing into the biopsychosocial effects of COVID on mental health, there has yet to be any analysis done between papers to determine not only the lasting effects of COVID, but also its effects on marginalized communities, self-diagnosis, and diagnostic criteria, which would provide the bigger picture on the broader reaches of COVID. This paper aims to examine the effects of this model on the relationship between COVID-19 and mental illnesses, specifically on aspects of the viral infection on physical and mental wellbeing, as well as the pandemic's interactions with social media and self-diagnosis.

Introduction

As societies recover from the devastating effects of COVID-19, beneath the rubble lies the unexpected result of an increased awareness around mental illnesses, especially amongst the younger generation. It is no secret that with the viral outbreak, many individuals suffered from the lasting health effects of COVID. This, as a result, may have caused tangible change within their mental health alongside their physical health. Additionally, the pandemic has uncovered the lack of support many marginalized communities encountered in regards to their mental health, signalling a point of reform needed to be made within the psychological and psychiatric field to address these needs. However, consequently, the pandemic also brought along the rising wave of social media through quarantine, leading to many adolescents self-diagnosing themselves with mental illnesses based on information gathered online. Self-diagnosis, as defined by Underhill and Foulkes (2024) is when individuals diagnose themselves with a mental health condition, either before or instead of seeing a professional. This can cause complications through
To complicate the situation even further, reforms continue to be made to the DSM-5-TR (Diagnostic and Statistical Manual of Mental Disorders) and ICD-11 (International Classification of Diseases), which has broadened the scope of mental illnesses, and the potential medicalization of everyday distress.

To investigate the relationship between these interconnecting factors, Endel's biopsychosocial model (1977) will be utilized to examine the weight of each argument, appraising evidence in support and against it, before finally evaluating its contributions to explaining the outcome of the entirety of COVID-19 on mental health between individuals. This essay will argue that COVID-19 caused corporeal change in mental health disorders and have enabled better recognition of mental health for marginalized communities, while also acknowledging that social media prevalence and the broadening of diagnostic categories may

have caused an influx of self-diagnoses that could become a concern within the near future, suggesting a need for a multifactorial perspective when addressing the issue of COVID-19 on mental illnesses.

The Corporeal Effects of COVID-19 on Mental Illnesses

Though COVID-19 is a viral infection that largely presents itself through biological symptoms and physical health conditions, it is not entirely separate from mental health. It may even have the capacity to act as a biological vulnerability prior to gaining a mental health diagnosis. Santomauro et al. (2021) conducted a large-scale study that found that there was a statistically significant increase in MDD and anxiety disorders worldwide after 2020, enough to show correlated patterns with increased infection rates. This suggests that COVID-19 had been attributed to real increases in mental illnesses through measuring change against pre-pandemic baselines. Findings of a large-scale study by Tauquet et al. (2021) further support this claim, with 33.6% of COVID-19 survivors receiving a neurological or psychiatric diagnosis within 6 months, of which anxiety disorders were most common. Additionally, the risks were higher after COVID-19 than the control group of demographically matched influenza survivors, concluding that there was substantial neurological and psychiatric morbidity after COVID-19, with the greatest risks in patients with severe COVID-19. This study can be modelled through the diathesis stress model, specifically the revised one proposed by Zuckerman (1999). This psychological model integrates psychological and biological factors to predict the likelihood of a disorder through a diathesis, also known as a vulnerability. In the context of Santomauro's study, COVID-19, the diathesis, when interacting with the surrounding social environment, can increase the risk of getting a mental disorder. While these studies support the argument that diagnoses increased because mental health genuinely worsened, both rely on correlational or observational data, meaning they cannot rule out other contributing factors such as heightened awareness or changes in help-seeking behavior. Even so, because these studies compare pre-pandemic baselines to matched illness groups, they make it difficult to reduce the rise in diagnoses to a case of self-diagnosis alone.

Furthermore, Haidt (2024) supports the idea that there are tangible reasons for worsening mental health and increased diagnoses. He argues that youth mental health has worsened as a result of a smartphone-based childhood, social media, and reduced independence. This "Great Rewiring of Childhood" is characteristic of a play-based childhood replaced with a phone-based childhood, contributing to a rise in anxiety, depression, and self harm from the early 2010s onward. As a result, adolescence has become too exposed to adult-like digital pressures and not enough to real-world experiences. Diagnosis may be rising because adolescence itself has become harder to navigate, causing young people to turn to social media which may have led to decreases in allowing them to have opportunities for free play and encounter uncertainty, core aspects of discovery, risks, and developing as an individual.

However, Haidt moved from correlation to causation too quickly, and did not consider the complexity of mental health trends in adolescents, though his argument still brings up important

aspects to consider when providing a holistic explanation. All of these arguments propose that the reason for an increase in diagnoses is due to a real change occurring, either as a result of the COVID pandemic or the rise of social media slowly replacing a play-based childhood. Haidt's work is persuasive in showing that some increases in diagnosis may reflect real declines in wellbeing rather than diagnostic inflation alone. Despite claims made by the author, this serves stronger as evidence that something has changed in youth mental health than as proof that smartphones are the main cause. Although many of these arguments dissuade from the idea of increased mental illness being intangible, they do not address other influences of social media that have cropped up in intensity as of late, namely self-diagnosis. From a biopsychosocial perspective, these findings represent the biological component of mental illness. COVID-19 functions not only as a viral infection but also as a biological stressor that interacts with psychological vulnerabilities and social environments, demonstrating that increases in diagnoses cannot be explained solely by changes in diagnostic practices.

The Self-Diagnosis Dilemma

Though it has been close to a decade since the pandemic first started, the effects of COVID-19 can still be felt today. It has affected public sectors, political climates, and most importantly: the mental health landscape. Within psychology, self-diagnosis, as defined by Underhill and Foulkes (2024) is when individuals diagnose themselves with a mental health condition, either before or instead of seeing a professional. This definition is important as it outlines the potential dangers surrounding self-diagnosis amongst individuals. For example, self-diagnosis could result in an incorrect or harmful diagnosis, increased anxiety surrounding diagnoses, or an attempt at self-treatment that could lead to the purchases of unprescripted medications (Monteith et al., 2024) The implications of this are that the distinction between genuine mental illnesses and ordinary psychological distress may be increasingly harder to navigate, making it more difficult to assess whether rising diagnoses reflect improved recognition or an increase in self-diagnosis.

Armstrong et al. (2025) stipulated that all participants had viewed mental health content in self-diagnosis online before entering treatment, with many using social media as a primary source of information over academically oriented websites. Participants who engaged more frequently with mental health content were more likely to place importance on obtaining a diagnosis. However, these findings do not establish whether social media is directly increasing the prevalence of mental illness or simply encouraging individuals to seek professional assessment by providing information that was previously inaccessible. Foster and Ellis (2024) built on this distinction, arguing that Tiktok-inspired self-diagnosis should not be reduced simply to misinformation or attention-seeking, but rather that self-diagnosis is shaped by identity, belonging, and self-labelling. Social media provides a framework through which adolescents can interpret distress, which is significant because it suggests that online platforms may spread both awareness and encouragement to adolescents to adopt diagnostic labels. Consequently, online platforms may simultaneously increase awareness of mental health conditions while encouraging

adoption of diagnostic labels. This creates further challenges when interpreting rising diagnosis rates, as increases may reflect both changes in help-seeking behavior and shifts in the diagnostic frameworks used by clinicians. Within Engel's biopsychosocial model, these findings illustrate the psychological dimension of mental illness. Social media influences how individuals perceive, interpret, and label their own symptoms, meaning that increased diagnoses may arise from additional changes in cognitive appraisal, identity formation, and help-seeking behavior. Therefore, to determine whether this growth in diagnoses represents improved recognition of mental illness or the medicalization of ordinary distress, it is also necessary to consider how revisions to diagnostic systems such as the DSM and ICD have influenced clinical practice.

In regards to the DSM, within Saving Normal, Allen Frances, a former chair of the DSM-IV task force, argued that modern psychiatry has expanded diagnostic categories too far. Frances (2013) acknowledges that psychiatric diagnosis is valuable when applied carefully, but argues that broader diagnostic criteria can risk medicalizing everyday life and create false epidemics, especially when considering the reach of influence the DSM has outside of psychology. For example, revisions made to the DSM have the potential to affect other areas of society such as political sectors or insurance companies. It can affect which individuals would be qualified for treatment, including the extent of which insurance will cover it, as well as their eligibility for benefits through welfare policies. From an education front, academic institutions rely on diagnostic criteria to justify special accommodations and behavioral interventions for students. Furthermore, diagnoses play a role in judicial fields through determining the extent of responsibility and punishment an individual should receive. From Engel's perspective, diagnosis is not merely a biological classification, but rather can be influenced by psychological and social contexts. Consequently, revisions to the DSM alter not only clinical practice, but also how individuals understand their own experiences and how society defines mental illness.

However, it could also be viewed that DSM-V have improved diagnostic specificity instead of solely broadening the categories. For example, the DSM-V has removed bereavement exclusions from MDD to better distinguish prolonged depressive symptoms from grief responses. This decision was based on findings of Zisook and Kendler (2008) which showed that the similarities outweighed the differences, concluding that bereavement-related major depression looked broadly similar to major depression associated with other stress life events. Therefore, bereavement-related depression was not shown to be clearly different enough to deserve automatic exclusion from a diagnosis. Therefore, while diagnostic criteria may contribute to over-diagnoses, its aim is not to medicalize everyday problems, but to adjust for better recognition of mental health conditions.

Discussion: The Systemic Mechanisms

The biopsychosocial model provides a useful framework for integrating these competing explanations. Rather than viewing increased diagnoses as solely the result of biological illness or diagnostic inflation, the model suggests that biological, cognitive, and social factors interact to shape mental health outcomes. For example, in context to how cultural norms, diagnostic

language, and social inequalities influence the experience and recognition of mental illness, there is evidence provided from ‘Saving Normal’ by Ethan Watters for an explanation. Watters (2010) argues that the way mental health is understood, diagnosed, and expressed is shaped by culture, and when Western psychiatric categories spread globally, they can change how people experience distress. In particular, he notes of an idea called ‘symptom pool’, meaning that each culture offers people a set of available ways to express suffering. People do not invent symptoms, rather drawing from the language, norms, and stories of their culture to identify symptoms. However, some increases in diagnosis may reflect the spread of diagnostic categories, specifically Western ones, which teach people how to recognize them in a Western clinical sense. For example, anorexia in Hong Kong was less focused on intense fear of fatness or body-image, rather more on bloatedness or stomach discomfort. However, after a highly publicized death and increased exposure to Western psychiatric models, anorexia started to resemble Western versions of it, increasing diagnoses based on Western clinical standards. This shows that diagnosis can shape the form suffering takes; through using Western diagnostic models it has shaped how distress and symptoms were viewed and as a result increased the number of diagnoses in a particular culture.

In the context of COVID-19 and social media, psychiatry professor Katherine Ramos (2022) stipulated that issues like loneliness, trauma, and social isolation became more visible as a result of COVID, but also argued that difficult emotions should be medicalized into a diagnosis, as most of it is a response to external factors. She proposed that while the pandemic may act as a traumatic stressor that could increase risk for mental illnesses and worsen pre-existing ones, mental health impacts need to be understood through an intersectional lens between interactions of age, race, ethnicity, disability, and socioeconomic status. To further support this, Arevian et al. (2020) proposes that the pandemic created major mental health consequences, exposing issues that were previously underaddressed and the disparities of the accessibility to services that lead to recovery. It highlighted that mental health was linked to concrete pressures that existed both as a result of COVID and even before it. This shows that the issue is not as simple as over-diagnosis, but rather that many underrepresented communities had finally been addressed as a result of COVID. When compared with Watters’ symptom pool, it shows how these underrepresented communities did not have the language to describe their issues, and it is only until after COVID that psychiatric diagnostic languages and models have finally expanded enough to include these communities. Though many of this surrounds a Western clinical model, it would account for the increase in diagnoses as better recognition. This suggests that broader psychiatric recognition can function as a way to correct historical neglect, particularly for marginalized groups. However, if every form of distress is translated into its own individual pathology, recognition risks turning into medicalization. Diagnosing every individual’s problems as symptoms of mental illnesses will lead them to face not only additional finances for treatments, but also stigma surrounding their diagnosis, which could cause further dissent and marginalization for these communities. Taken together, these sources suggest that rising

diagnoses may reflect the cultural spread of Western psychiatric language as well as the overdue recognition of distress in groups that were historically overlooked.

Conclusion

Ultimately, mental illness today can not be simply defined as over-diagnosed or better recognized. The evidence suggests that genuine increases in psychological distress, especially in response to COVID-19 and major societal changes, have led to more accurate recognition of mental health difficulties in many individuals and communities. However, at the same time, the rise of self-diagnosis within social media and the expansion of diagnostic criteria has created a risk that ordinary distress may increasingly be interpreted through psychiatric labels. This means that the issue has now turned to how to distinguish between serious mental illness as responses to adversity without ignoring marginalized individuals who did not have a voice prior to this. Ultimately, the biopsychosocial model best explains why no single explanation adequately accounts for the rise in mental illness diagnoses after COVID-19. Biological consequences of infection, psychological processes of self-diagnosis, and social influences of cultural and unequal access to care interact to produce the observed increase in diagnoses.

Works Cited

Arevian, Armen C., et al. "Mental Health Community and Health System Issues in COVID-19: Lessons from Academic, Community, Provider and Policy Stakeholders." Ethnicity & Disease, vol. 30, no. 4, 2020, pp. 695–700. JSTOR, https://www.jstor.org/stable/48668042. Accessed 5 June 2026.

Armstrong, S., Osuch, E., Wammes, M., Chevalier, O., Kieffer, S., Meddaoui, M., & Rice, L. (2025). Self-diagnosis in the age of social media: A pilot study of youth entering mental health treatment for mood and anxiety disorders. Acta Psychologica, 256, 105015. https://doi.org/10.1016/j.actpsy.2025.105015

Frances, A. (n.d.). Saving normal : an insider's revolt against out-of-control psychiatric diagnosis, dsm-5, big pharma, and the medicalization of ordinary life.

Haidt, J. (n.d.). The anxious generation : how the great rewiring of childhood is causing an epidemic of mental illness.

Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E. D., & Bauer, M. (2024). Implications of online self-diagnosis in psychiatry. Pharmacopsychiatry, 57(02), 45-52. https://doi.org/10.1055/a-2268-5441

Ramos, Katherine. "Mental Health Impacts of the COVID-19 Pandemic." Generations: Journal of the American Society on Aging, vol. 46, no. 1, 2022, pp. 1–8. JSTOR, https://www.jstor.org/stable/48679953. Accessed 5 June 2026.

Santomauro, D. F., Mantilla Herrera, A. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., Abbafati, C., Adolph, C., Amlag, J. O., Aravkin, A. Y., Bang-Jensen, B. L., Bertolacci, G. J., Bloom, S. S., Castellano, R., Castro, E., Chakrabarti, S., Chattopadhyay, J., Cogen, R. M., Collins, J. K., . . . Giles, J. R. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic. The Lancet, 398(10312), 1700-1712. https://doi.org/10.1016/s0140-6736(21)02143-7

Taquet, M., Geddes, J. R., Husain, M., Luciano, S., & Harrison, P. J. (2021). 6-month neurological and psychiatric outcomes in 236 379 survivors of covid-19: A retrospective cohort study using electronic health records. The Lancet Psychiatry, 8(5), 416-427. https://doi.org/10.1016/s2215-0366(21)00084-5

Turuba, R., Cormier, W., Zimmerman, R., Ow, N., Zenone, M., Quintana, Y., Jenkins, E., Ben-David, S., Raimundo, A., Marcon, A. R., Mathias, S., Henderson, J., & Barbic, S. (2024). Exploring how youth use tiktok for mental health information in British Columbia: Semistructured interview study with youth. JMIR Infodemiology, 4, e53233. https://doi.org/10.2196/53233

Underhill, R., & Foulkes, L. (2025). Self-Diagnosis of Mental Disorders: A Qualitative Study of Attitudes on Reddit. Qualitative health research, 35(7), 779–792. https://doi.org/10.1177/10497323241288785

Watters, E. (n.d.). Crazy like us : the globalization of the American psyche.

Coupled Multi-Body Dynamics and Aerodynamic Load Modeling for Predicting Limit Handling in High-Downforce Hypercars By Aanya Garg

Abstract

Predicting limit handling in high-downforce hypercars is a complex engineering challenge due to the intricate coupling of aerodynamic forces and multi-body dynamics (MBD). This study investigates the current methodologies, computational bottlenecks, and suspension strategies utilized in the industry through a mixed-methods approach, combining a targeted survey of 33 industry experts with a comprehensive literature review. The findings reveal that pitch angle variations pose the highest risk for inducing sudden aerodynamic instability, such as diffuser vortex breakdown, making the implementation of decoupled suspension topologies essential. Furthermore, while fully coupled two-way computational fluid dynamics (CFD) and MBD co-simulations offer significantly higher accuracy for transient maneuvers, their widespread adoption is severely hindered by extreme computational costs and dynamic fluid meshing difficulties. As a result, mid-fidelity simulation tools utilizing Vortex Particle Methods (VPM) are emerging as highly effective and efficient alternatives. Finally, the study highlights a unanimous industry consensus that standard empirical tire models are inadequate at the aerodynamic limit due to severe tire load sensitivity, necessitating the transition to advanced structural tire models. These insights provide a clear roadmap for optimizing vehicle simulation architectures in high-performance automotive engineering.

Introduction

The engineering of high-performance vehicles has undergone a profound paradigm shift over the last several decades, transitioning from an era dominated by purely mechanical optimization to one defined by highly integrated multidisciplinary systems. Historically, vehicle system dynamics emerged from the analysis of rigid body mechanics, utilizing mechanical multi-body system (MBS) models to simulate and predict a vehicle’s longitudinal, lateral, and vertical behavior (Arnold et al., 2007). In this traditional framework, aerodynamic forces were largely treated as parasitic penalties, static drag vectors that required mitigation to achieve higher top speeds (MPH, 2024). However, the relentless pursuit of performance in motorsport and the subsequent development of modern hypercars have fundamentally inverted this relationship. For high-downforce vehicles operating at extreme velocities, the fluid dynamics of the air passing over, under, and around the vehicle no longer represent mere resistance; rather, the air acts as a critical structural force that dictates the absolute boundaries of the vehicle's dynamic capabilities (MPH, 2024).

This evolution has elevated aerodynamic downforce to a primary performance discriminator. As a vehicle's longitudinal speed increases, aerodynamic surfaces,most notably the underbody floor, rear diffusers, and inverted wings,generate massive vertical loads (Marchesin et al., 2017; MPH, 2024). These loads, which often exceed the static weight of the vehicle itself, artificially press the tires into the track surface, fundamentally altering the normal forces exerted

on the tire contact patches without adding significant inertial mass (Marchesin et al., 2017). This phenomenon, colloquially termed "aerodynamic grip," operates in parallel with the mechanical grip derived from the vehicle's suspension geometry, mass distribution, and tire compounds (Milliken & Milliken, 1995; Continental, 2024). By exponentially increasing the vertical load on the tires, aerodynamic downforce vastly expands the vehicle's traction circle, permitting unprecedented cornering speeds, braking decelerations, and lateral accelerations that would be physically impossible under the constraints of static weight alone (Milliken & Milliken, 1995; Continental, 2024).

However, the generation of this aerodynamic force is exceptionally non-linear and exquisitely sensitive to the vehicle's instantaneous spatial attitude (Carreira, 2020; MDPI, 2026). As a high-downforce hypercar navigates a circuit, it is subjected to continuous and violent variations in pitch, roll, yaw, and heave (Carreira, 2020). These variations are driven by driver inputs, track surface irregularities, and inertial load transfers during transient maneuvers (Carreira, 2020; Hendy, 2019). Because ground-effect aerodynamics rely on the precise proximity of underbody devices to the track surface, these attitude changes continuously alter the efficiency of the aerodynamic surfaces, creating a highly transient and volatile aerodynamic state (Marchesin et al., 2017). Consequently, the aerodynamic forces and pitching moments applied to the multi-body chassis are in a state of constant flux. This creates a complex, closed-loop feedback system: the vehicle's aerodynamic performance directly alters its mechanical kinematics by compressing the suspension, while the resulting kinematic displacement simultaneously alters the vehicle's aerodynamic performance (Marchesin et al., 2017). Understanding, simulating, and predicting the limit handling behavior of such vehicles,defined as the vehicle's transient response when operating at the absolute peak of its tire adhesion capabilities,requires sophisticated computational paradigms (Pacejka, 2012). Limit handling represents the fragile boundary where the linear elastic behavior of the tire transitions into non-linear sliding friction. When a vehicle operates at this limit under extreme aerodynamic loads, it resides in a highly sensitive region of the tire's friction curve (Milliken & Milliken, 1995). A sudden loss of downforce, triggered by an abrupt change in ride height, a diffuser stall, or a severe crosswind gust, causes an instantaneous and catastrophic reduction in the tire's lateral force capability (Marchesin et al., 2017; Milliken & Milliken, 1995). This sudden degradation leads to rapid, unrecoverable breakaway, making the prediction of limit handling an essential component of vehicle safety and performance optimization.

Traditional quasi-steady-state assumptions, which have historically formed the backbone of vehicle simulation, are entirely insufficient for capturing these highly transient phenomena (SAE, 2016; SAE, 2017). Quasi-steady models treat aerodynamic loads as static forces mapped solely to vehicle velocity, failing to account for the dynamic interactions between the suspension's non-linear stiffness and the fluid boundary layers (Marchesin et al., 2017; SAE, 2016). To accurately capture the nuances of high-downforce limit handling, the automotive and motorsport industries have shifted toward the co-simulation of Multi-Body Dynamics (MBD) and Computational Fluid Dynamics (CFD) (Arnold et al., 2007; PreCICE, 2021).

This coupled, multi-physics approach allows for the simultaneous calculation of structural kinematics and fluid dynamics (PreCICE, 2021; Hexagon, 2024). It ensures that the aerodynamic loads applied to the vehicle accurately reflect its instantaneous spatial attitude and, conversely, that the vehicle's attitude responds accurately to the instantaneous aerodynamic forces generated by complex fluid behaviors such as wake hysteresis and vortex breakdown (PreCICE, 2021). Co-simulation bridges the gap between mechanical engineering and aerodynamics, providing a comprehensive digital twin capable of predicting pitch instability, crosswind deviation, and aero-balance migration before a physical prototype is ever manufactured (Marchesin et al., 2017; SAE, 2017).

This comprehensive research report provides an exhaustive, meticulous analysis of the coupling between multi-body vehicle dynamics and aerodynamic load modeling. The subsequent sections will systematically explore the foundational theories of high-downforce fluid dynamics, the complex viscoelastic mechanics of tire load sensitivity, the necessity of specialized suspension topologies such as third springs and non-linear bump stops, and the advanced numerical co-simulation methodologies required to accurately predict limit handling in modern aerodynamic hypercars. Through an in-depth examination of the latest literature, computational methodologies, and dynamic modeling techniques, this report establishes a complete theoretical and practical framework for understanding the aero-mechanical coupling that dictates extreme vehicle performance.

Fundamental Principles of High-Downforce Aerodynamics

At speeds exceeding 200 mph, the atmosphere becomes a dense, structural medium that hypercars must manipulate to maintain stability and grip (MPH, 2024). The primary objective of aerodynamic design in this context is to negotiate the inherent tradeoff between drag (air resistance that impedes forward velocity) and downforce (vertical pressure that pushes the vehicle into the track) (MPH, 2024).

The Generation of Aerodynamic Forces

The aerodynamic forces acting upon a multi-body vehicle system are mathematically defined by the dynamic pressure of the freestream air, the vehicle's frontal reference area, and dimensionless aerodynamic coefficients derived from the vehicle's specific geometry (Carreira, 2020; Occam's Racers, 2019). These forces are expressed through the following fundamental fluid dynamics equations:

F_z = \frac{1}{2} \rho v^2 A C_L \\ F_x = \frac{1}{2} \rho v^2 A C_D \\ M_y = \frac{1}{2} \rho v^2 A L C_M

Where F_z is the vertical downforce, F_x is the longitudinal drag, M_y is the pitching moment, \rho represents air density, v is the vehicle velocity, A is the reference area, L is the characteristic length (often wheelbase), and C_L, C_D, and C_M represent the respective

coefficients for lift, drag, and pitch (Carreira, 2020; Occams Racers, 2019; Marchesin et al., 2017). In high-performance applications, the lift coefficient C_L is typically a negative value, indicating downforce generation (Occams Racers, 2019).

While early aerodynamic implementations relied heavily on prominent, aggressive front and rear wings, these devices incur severe drag penalties due to profile drag and induced drag from wingtip vortices (Pacejka, 2012; MPH, 2024). Consequently, modern hypercars derive the vast majority of their vertical load from underbody airflow management, utilizing flat floors and rear diffusers (MPH, 2024). A flat underbody panel smooths the turbulent air passing beneath the chassis, accelerating the flow and dropping the static pressure in accordance with Bernoulli's principle (MPH, 2024). The rear diffuser subsequently acts as an expansion chamber, aggressively accelerating the airflow underneath the car to match the freestream velocity at the rear, thereby creating a massive low-pressure zone that effectively vacuums the vehicle onto the track (Marchesin et al., 2017; MPH, 2024).

Ground Effect and Ride Height Sensitivity

The defining characteristic of these underbody aerodynamic devices is their extreme sensitivity to ground clearance, a physical phenomenon known as the "ground effect" (Marchesin et al., 2017). The ground effect dictates that as an aerofoil, flat floor, or diffuser moves closer to the track surface, the downforce coefficient increases in a highly non-linear fashion (Marchesin et al., 2017). This amplification is primarily attributed to the restriction of fluid flow between the underbody and the ground, which forces a more dramatic acceleration of the air, as well as the physical interruption of tip vortices that would otherwise bleed higher-pressure ambient air into the critical low-pressure underbody region (Marchesin et al., 2017).

Because of the ground effect, the aerodynamic performance of a hypercar is intrinsically linked to its instantaneous ride height (Marchesin et al., 2017). As vehicle speed increases, the generated downforce compresses the suspension springs, which lowers the ride height (Marchesin et al., 2017). This lower ride height further enhances the downforce coefficient, creating a positive feedback loop that continues until the suspension's mechanical limits (bump stops) are reached or a fluid dynamic stall occurs (Marchesin et al., 2017). Understanding this unsteady aerodynamic behavior during transient maneuvers requires analyzing three major effects: ground clearance (representing aerodynamic stiffness), induced angle of attack (representing aerodynamic damping), and added mass (representing aerodynamic inertia, though this is often neglected for road holding frequencies) (Marchesin et al., 2017).

Wheel-Ground Conditions and Attitude Modulation

The aerodynamic flow field around a hypercar is further complicated by the interaction between the rapidly rotating wheels, the stationary ground, and the moving vehicle body (MDPI, 2026). Recent computational numerical simulations and wind tunnel validations have demonstrated that the specific boundary conditions of the wheel-ground interface profoundly impact the total vehicle aerodynamics (MDPI, 2026).

To isolate these effects, researchers evaluate three distinct wheel-ground operational combinations:

1. MR (Moving ground with Rotating wheels): Represents the true physical state of a vehicle in motion on a circuit.
2. MS (Moving ground with Stationary wheels): Utilized to isolate the effects of ground boundary layers without the complex rotational wake generated by the tire treads.
3. SS (Stationary ground with Stationary wheels): The most simplified and historically common wind tunnel condition (MDPI, 2026).

Research reveals a clear dichotomy in how these conditions affect the vehicle's aerodynamic forces. Wheel conditions,specifically the rotation of the rear wheels and their localized wake interactions with the diffuser-equipped body,predominantly dictate the vehicle's overall aerodynamic drag (MDPI, 2026). In digital simulations, the contact patch between the tire and the ground must be modeled as a small convex feature to avoid unphysical, abrupt meshing connections that artificially corrupt this wake (MDPI, 2026).

Conversely, ground conditions primarily influence the aerodynamic downforce generated by the main vehicle body (excluding the wheels) and the front wing (MDPI, 2026). The downforce variations caused by moving ground boundary layers subsequently alter the induced drag of the vehicle, which further amplifies differences in total drag (MDPI, 2026).
Crucially, the impact of these ground conditions on the front wing is directly modulated by the vehicle's attitude (MDPI, 2026). Depending on the pitch and roll angles of the chassis, these ground interactions can result in either an amplification or a severe degradation of front wing downforce (MDPI, 2026). Because the front wing is highly sensitive to these attitude-ground interactions, changes in pitch instantly shift the aerodynamic force distribution between the front and rear axles,altering the car's aerodynamic balance (aerobalance) and directly impacting handling stability (MDPI, 2026).

Tire Dynamics and Load Sensitivity at the Limit

The ultimate objective of managing complex aerodynamic flows is to increase the vertical load (F_z) on the vehicle's tires, thereby expanding the vehicle's friction limits and increasing its limit handling capabilities (Milliken & Milliken, 1995). However, the relationship between this vertical aerodynamic load and a tire's ability to generate lateral (F_y) and longitudinal (F_x) forces is highly non-linear, requiring sophisticated tire models to bridge the gap between aerodynamics and multi-body dynamics.

Classical Friction Theory vs. Viscoelastic Tire Behavior

Classical Coulomb friction theory postulates that the maximum frictional force developed between two surfaces is directly proportional to the normal load applied (F_f = \mu F_N). This implies that the coefficient of friction (\mu) remains entirely constant regardless of the vertical

load applied (Milliken & Milliken, 1995). Furthermore, according to Amonton's First and Second Laws of friction, the force of friction is independent of the contact area, and Coulomb's Law dictates it is independent of velocity (Occam's Racers, 2023).

If racing tires adhered to classical rigid-body physics, aerodynamic downforce would yield a perfectly linear, infinitely scalable increase in cornering capability, and narrow tires would perform identically to wide tires (Occams Racers, 2023). In reality, conventional pneumatic tires are constructed from complex viscoelastic rubber compounds, and their frictional properties deviate significantly from classical macroscopic models (Milliken & Milliken, 1995). Grip is generated through two primary microscopic mechanisms: adhesion (molecular bonding between the rubber and the track) and hysteresis (energy loss as the rubber deforms around the macro-texture of the road surface). Both of these mechanisms saturate under extreme loads.

The Non-Linearity of Tire Load Sensitivity

Tire load sensitivity describes the critical empirical phenomenon wherein the coefficient of friction of a tire actually *decreases* as the vertical load increases (Milliken & Milliken, 1995). While the absolute maximum lateral force (F_y) does increase with added vertical load,which explains the immense cornering speeds of aerodynamic hypercars,it does so at a rapidly diminishing rate (Milliken & Milliken, 1995).

Mathematically, the maximum horizontal force generated by the tire is roughly proportional to the vertical load raised to a fractional exponent:

F_{y, max} \propto F_z^n

Where the load sensitivity exponent n typically ranges between 0.7 and 0.9 for high-performance racing tires (Milliken & Milliken, 1995).

This non-linear sensitivity is exhaustively documented in vehicle dynamics literature. For instance, data extracted from Milliken and Milliken's fundamental text *Race Car Vehicle Dynamics* illustrates how peak friction degrades. As vertical load increases from 900 lbf to 1800 lbf, the peak lateral friction coefficient (F_y/F_z) drops from 1.10 to 0.97 (Milliken & Milliken, 1995). Furthermore, the slip angle (the angle between the tire's actual direction of travel and the direction it is pointing) required to achieve this peak force increases from 5.6 degrees to 6.7 degrees (Milliken & Milliken, 1995).

Vertical Load (F_z, lbf)	Peak Friction Coefficient (F_y / F_z)	Slip Angle for Peak Force (Degrees)
900	1.10	5.6
1350	1.08	6.0
1800	0.97	6.7
Table 1: Example of Tire Load Sensitivity demonstrating the		

Vertical Load (F_z, lbf)	Peak Friction Coefficient (F_y / F_z)	Slip Angle for Peak Force (Degrees)
degradation of peak friction coefficient and increasing slip angles under increasing vertical loads (Milliken & Milliken, 1995).		

To accurately model these effects within multi-body simulations, engineers rely on highly parameterized empirical mathematical models, the most prominent being the Pacejka "Magic Formula" (Pacejka, 2012). The Magic Formula calculates longitudinal and lateral forces based on vertical load, slip angle, slip ratio, and camber, utilizing complex polynomial regressions to capture the non-linear degradation of grip at extreme aerodynamic loads (Rutgers, 2024; Pacejka, 2012).

Implications for Lateral Weight Transfer and Limit Handling

The load sensitivity of tires fundamentally dictates a vehicle's limit handling behavior, particularly during lateral load transfer (body roll). When a hypercar corners, inertial forces act on the center of gravity, causing weight to transfer from the inside tires to the outside tires (Pacejka, 2012). Because of tire load sensitivity, the gain in grip on the heavily loaded outside tire is substantially less than the simultaneous loss of grip on the lightly loaded inside tire (Marchesin et al., 2017; Pacejka, 2012). Therefore, the total combined lateral force capability of the axle decreases as load transfer increases (Marchesin et al., 2017; Pacejka, 2012).

Aerodynamic downforce mitigates this physical limitation by globally increasing the vertical load on all four tires without adding significant inertial mass or raising the vehicle's center of gravity (Marchesin et al., 2017; Pacejka, 2012). By artificially pressing the vehicle into the track, downforce increases the absolute lateral force capability, allowing for much higher cornering speeds before the slip angles breach the limit of adhesion (Continental, 2024).

However, because hypercar tires operate under these extreme aerodynamic loads, they reside in a highly sensitive, saturated region of their friction curves. If a sudden loss of downforce occurs,such as a diffuser stall due to an abrupt change in ride height, or a disruption in flow caused by a severe crosswind gust,it causes an instantaneous and catastrophic reduction in the tire's force capability. This sudden contraction of the tire's traction circle leads to immediate, unrecoverable breakaway at the handling limit, highlighting why precise aerodynamic platform control is mandatory (Marchesin et al., 2017; Milliken & Milliken, 1995).

Multi-Body Dynamics (MBD) and Kinematic Modeling

Multi-Body Dynamics (MBD) provides the rigorous mathematical framework required to simulate the mechanical response of a vehicle to internal geometric constraints and external aerodynamic forces (Arnold et al., 2007). In MBD simulations, a hypercar is represented as a

complex assembly of rigid and flexible bodies interconnected by kinematic joints, compliant force elements (springs, dampers, anti-roll bars), and the tire-road interface model (PreCICE, 2021).

Mathematical Formulation of MBD

The state of an MBD system is continuously calculated by solving an extensive system of Differential-Algebraic Equations (DAEs) (MBDyn, 2024). The Newton-Euler equations describe the translational and rotational motion of the vehicle's unsprung and sprung masses, while Lagrange multipliers are utilized to mathematically enforce kinematic constraints, such as the exact pivot points of suspension wishbones and steering tie rods (MBDyn, 2024).
The extended equations of motion for a coupled multi-body vehicle system can be represented as:

$$M(q)\ddot{q} + k(q, \dot{q}) + C_q^T \lambda = Q_e + Q_a(q, \dot{q}, t)$$

Where $M(q)$ is the mass matrix of the system, q represents the generalized coordinates of the bodies, $k(q, \dot{q})$ contains the Coriolis, centrifugal, and gyroscopic forces, $C_q^T \lambda$ represents the internal constraint forces via the Jacobian matrix C_q and Lagrange multipliers λ, Q_e denotes external mechanical forces (such as tire contact patch forces), and $Q_a(q, \dot{q}, t)$ represents the highly transient aerodynamic loads (Arnold et al., 2007; MBDyn, 2024).

Because of the extreme stiffness of hypercar suspension components and the rapid onset of aerodynamic forces, these differential equations become highly "stiff" in a mathematical sense. Solvers such as MBDyn, SIMPACK, and specialized MATLAB/Simulink co-simulations must use advanced multi-step, L-stable integration algorithms with algorithmic dissipation to numerically evaluate these equations in the time domain without diverging into numerical instability (Arnold et al., 2007; PreCICE, 2021; MBDyn, 2024).

Suspension Topologies for Aerodynamic Platforms

The integration of massive aerodynamic forces fundamentally conflicts with the traditional objectives of suspension tuning. In a conventional passenger vehicle, or a racecar devoid of aerodynamic aids, suspension springs and dampers are tuned primarily to isolate the sprung mass from road irregularities. This compliance maximizes mechanical grip by ensuring the tire contact patch load remains as consistent as possible over bumps (Marchesin et al., 2017; Racecar Engineering, 2020).

However, high-downforce hypercars generate aerodynamic loads that would instantly crush a compliant suspension, forcing the chassis to bottom out against the track surface in a shower of sparks (Racecar Engineering, 2020). Furthermore, because the vehicle's aerodynamic map is exquisitely sensitive to millimeter changes in ride height, pitch, and roll, the suspension must hold the vehicle platform in a rigid, highly specific operating window to prevent catastrophic aero-stall (Marchesin et al., 2017; Hendy, 2019).

Chassis Modes and the Heave Spring (Third Spring)

To resolve the contradictory requirements of a rigid aerodynamic platform and a compliant mechanical footprint, modern hypercars and aerodynamic racecars decompose suspension movement into four distinct chassis displacement modes: Heave, Pitch, Roll, and Warp (Racecar Engineering, 2020).

Chassis Mode	Description	Primary Influence
Heave (Bounce)	Uniform vertical translation of the chassis. Both wheels on an axle compress or extend simultaneously.	Ride height control, ground-effect downforce generation, diffuser clearance, and aerodynamic stall prevention (Racecar Engineering, 2020; Diaz, 2019).
Pitch	Rotation about the lateral axis (e.g., diving under braking, squatting under acceleration).	Aerodynamic balance shift (center of pressure migration) and longitudinal weight transfer (Hendy, 2019).
Roll	Rotation about the longitudinal axis during cornering, pivoting around the geometric roll center.	Lateral load transfer, tire camber control, and mechanical grip optimization (Racecar Engineering, 2020; Ortiz, 2014).
Warp	Cross-axle articulation (e.g., front left and rear right compress, while the others extend).	Navigating track undulations and kerbs without unsettling the platform (Racecar Engineering, 2020).
Table 2: Fundamental vehicle chassis displacement modes and their influence on hypercar dynamics (Racecar Engineering, 2020; Diaz, 2019).		

To independently control these modes, engineers utilize complex suspension topologies featuring a "third spring" or "heave spring" (OptimumG, 2021; Wavey Dynamics, 2022). In this arrangement, the conventional left and right corner springs (which primarily handle roll) are made relatively soft to allow for compliant roll dynamics, maximizing mechanical grip during cornering (Ortiz, 2014; Diaz, 2019).

However, when aerodynamic downforce pushes the entire chassis downward simultaneously (heave mode), a mechanical linkage actuates the heave spring (Diaz, 2019; Wavey Dynamics, 2022). This third spring is exceptionally stiff,often operating at natural frequencies between 3.0 Hz and 7.0 Hz for aerodynamic racecars (OptimumG, 2021). The heave spring provides the massive vertical stiffness required to support aerodynamic loads at high speeds and maintain the vehicle's critical ride height without unnecessarily stiffening the car in roll (Ortiz, 2014; Wavey Dynamics, 2022). This specific decoupling of roll and heave vasty increases the performance window of the vehicle, allowing it to navigate slow, mechanically demanding corners while surviving high-speed, aerodynamically punishing straights (Diaz, 2019; Wavey Dynamics, 2022).

Non-Linear Stiffness and Bump Stops

In addition to heavy springs, the vertical dynamics of high-downforce vehicles rely heavily on non-linear stiffness devices known as bump stops (or packers) (Marchesin et al., 2017). A bump stop is an elastomeric or polyurethane element that operates in parallel with the main suspension springs, but it only engages after a specific amount of suspension travel,defined as the bump stop gap,has been consumed (Marchesin et al., 2017; OptimumG, 2019).
As the vehicle accelerates and downforce increases proportionally to the square of velocity, the suspension compresses through the "soft" initial spring gap until the dampers contact the bump stops (Marchesin et al., 2017; OptimumG, 2019). At the exact moment of engagement, the overall suspension stiffness increases exponentially, effectively freezing the vehicle's ride height at its optimal aerodynamic trim (OptimumG, 2019). Multi-body simulations must meticulously model the precise gap and non-linear stiffness curves of these bump stops to predict when they engage at specific velocities, as this sudden engagement drastically alters the vehicle's frequency response, damping ratios, and load transfer characteristics (Marchesin et al., 2017; OptimumG, 2019).

The Aero-Kinematic Coupling and Dynamic Instabilities

The true complexity of limit handling prediction lies in the coupled nature of vertical dynamics and unsteady aerodynamics (Marchesin et al., 2017). A hypercar in motion does not simply achieve a steady state; it is a dynamic oscillator constantly perturbed by road inputs, driver steering actions, and transient fluid dynamics (Marchesin et al., 2017).

Aerodynamic Mapping and Surrogate Modeling

To feed accurate aerodynamic loads into MBD models without incurring the computational impossibility of running full Navier-Stokes CFD simulations at every millisecond time step, engineers utilize Aerodynamic Maps (Aero Maps) (Hendy, 2019; Monash, 2024; Catesby, 2024). An aero map is a multi-dimensional surrogate model,a complex, multi-variable lookup table,that quantifies the vehicle's aerodynamic coefficients (C_L, C_D, C_M) as a function of its specific spatial attitude (Carreira, 2020; Hendy, 2019; Catesby, 2024).

A comprehensive aero map is parameterized by multiple degrees of freedom defining the vehicle's state:

- Front Ride Height (frh): Distance from the front axle reference point to the ground, dominating front wing and splitter effectiveness (Hendy, 2019).
- Rear Ride Height (rrh): Distance from the rear axle reference point to the ground, critical for diffuser expansion ratios (Hendy, 2019).
- Roll Angle: Body rotation that creates asymmetrical underbody flow and can unseal the floor edges (Hendy, 2019).
- Yaw Angle: The slip angle of the vehicle relative to freestream air, critical for assessing crosswind stability (Hendy, 2019).

- Steer Angle: The deflection of the front wheels, which drastically disrupts the wake interacting with sidepods and the leading edge of the underfloor (Hendy, 2019).

Generating these extensive maps requires executing hundreds, or even thousands, of discrete CFD simulations across a matrix of vehicle attitudes (Carreira, 2020). Advanced computational meshing techniques, such as hexcore volumes and automated kinematic ride height changers, are used to automatically morph or reposition the vehicle geometry for each individual data point in the matrix (Hendy, 2019; Catesby, 2024). The sensitivity revealed by these maps can be extreme. For example, a detailed CFD study conducted on the FST10e prototype demonstrated that yaw angle was the most impactful sensitivity parameter, resulting in a staggering 16% change in total downforce strictly across the standard working interval of slip angles (Carreira, 2020).

Once generated, this multi-dimensional aero map is imported into the MBD solver (such as OptimumDynamics, SIMPACK, or ADAMS) (OptimumG, 2019; PreCICE, 2021). At every integration time step of the simulation, the MBD solver queries the aero map based on the vehicle's instantaneous kinematics, interpolates the exact aerodynamic forces, and applies them to the chassis center of pressure to calculate the subsequent mechanical displacement (Hendy, 2019; OptimumG, 2019).

Pitch Instability, Aerobalance Migration, and Diffuser Stall

The rigid coupling of suspension and aerodynamics frequently exposes critical vulnerabilities in vehicle limit handling, the most notable and dangerous being the phenomenon of pitch instability (Marchesin et al., 2017).

While high levels of downforce decrease lap times, the longitudinal position of the aerodynamic center of pressure migrates dynamically with the pitch attitude of the vehicle (Marchesin et al., 2017; MPH, 2026). If a hypercar dives sharply under heavy braking, the front ride height decreases while the rear ride height simultaneously increases. This attitude shifts the aerodynamic balance sharply to the front axle. This aero-migration overloads the front tires while simultaneously starving the rear tires of vertical load, causing the rear end to lose lateral capacity and inducing immediate, snap oversteer at corner entry (Hendy, 2019; MDPI, 2022).

More critically, at sustained high speeds (e.g., between 100 km/h and 150 km/h for certain high-downforce formulas), the massive aerodynamic pressure compresses the rear suspension to the point where the rear ride height drops below a critical fluid-dynamic threshold (typically around 20 mm) (Marchesin et al., 2017). At this extreme proximity to the ground, the boundary layer of the airflow passing through the diffuser chokes, leading to sudden, catastrophic diffuser stall (Marchesin et al., 2017).

When the diffuser stalls, rear downforce drops precipitously. The sudden loss of aerodynamic load instantly unloads the rear suspension, causing the rear of the car to spring upward on its stiff springs. As the ride height increases, the diffuser clears the boundary layer and "unstalls," immediately generating massive downforce again, sucking the car violently back

down to the track (Marchesin et al., 2017). This cycle triggers a violent, undamped aerodynamic oscillation known in motorsport as "porpoising" or bouncing (Marchesin et al., 2017).
MBD stability analysis reveals that during these high-speed regimes, the front-end and rear-end bounce modes can exhibit negative damping values, meaning the system's eigenvalues migrate into the unstable positive real axis of the complex plane (Marchesin et al., 2017). In these scenarios, any slight perturbation from the road surface triggers vertical vibrations that grow continuously in magnitude, destroying tire contact patch load consistency and completely destabilizing the vehicle at the limit (Marchesin et al., 2017). Remedial tuning requires altering the mechanical suspension,utilizing softer bump stops with larger initial gaps, paired with highly specific, non-linear damper curves,to prevent the vehicle from breaching the aerodynamic stall threshold, even if it sacrifices peak theoretical downforce (Marchesin et al., 2017; OptimumG, 2019).

Co-Simulation Architectures: Bridging CFD and MBD

While multi-dimensional aerodynamic surrogate maps provide excellent approximations for quasi-steady states and broad attitude sweeps, limit handling often involves highly transient fluid dynamics,such as wake hysteresis, sudden crosswind gusts, separated flows, and vortex breakdown,that simply cannot be captured by static lookup tables (SAE, 2016; PreCICE, 2021). To capture the true temporal physics of limit handling, engineers must utilize advanced co-simulation architectures that explicitly couple the MBD and CFD solvers in the time domain (Arnold et al., 2007; PreCICE, 2021).

Partitioned Co-Simulation and Data Exchange Methods

Because the Navier-Stokes equations governing fluid dynamics and the Differential-Algebraic Equations governing multi-body kinematics are inherently different mathematical structures requiring vastly different temporal resolutions, monolithic solvers (which attempt to solve both physics simultaneously in a single immense matrix) are computationally prohibitive (PreCICE, 2021). Instead, partitioned co-simulation approaches are deployed (PreCICE, 2021).

In a partitioned scheme, the fluid and structural domains are solved by dedicated, specialized software packages (for example, MBDyn handling the multibody dynamics, coupled via a multi-physics library like preCICE to a fluid solver like OpenFOAM or DUST) (PreCICE, 2021; MBDyn, 2024). The solvers execute their time integrations independently but pause at predefined, strict communication intervals to exchange boundary conditions (PreCICE, 2021). The MBD solver sends the instantaneous position, velocity, and orientation of the vehicle body and wheels to the CFD solver. The CFD solver subsequently updates its fluid mesh, calculates the resultant fluid pressure and shear stresses, integrates these to find the aerodynamic forces and moments, and passes these highly accurate transient loads back to the MBD solver (PreCICE, 2021; SAE, 2016).

Careful selection of the communication interval is paramount; explicit loose coupling

schemes can lead to severe numerical instability, particularly in systems with closed kinematic loops, if the data exchange frequency is not tuned correctly (PreCICE, 2021). Typical high-fidelity vehicle co-simulations operate with sampling frequencies around 200 Hz to maintain stability while capturing critical dynamic frequencies (PreCICE, 2021).

6.2 One-Way vs. Fully Coupled (Two-Way) Simulation

The fidelity of the co-simulation strategy heavily dictates its capability to predict limit handling accurately (SAE, 2016).

One-Way Transient Coupling: In one-way coupling, a CFD simulation is executed to drive the vehicle through a prescribed aerodynamic environment, such as a time-dependent crosswind gust (SAE, 2017). The CFD predicts the transient aerodynamic forces as the gust hits the vehicle, but it does so without updating the vehicle's spatial attitude in response to the wind (SAE, 2017). These unfiltered force histories are then imported into the MBD simulation as external applied forces to evaluate the vehicle's subsequent yaw, pitch, and roll response (SAE, 2017).

This methodology is frequently used in the automotive industry to mimic open-loop test procedures, such as ISO 12021:2010 (Sensitivity to lateral wind) (SAE, 2016; SAE, 2017). Studies conducted on vehicles like the Ford Escort, Ford Everest, and Ford B-Max demonstrated that one-way transient coupling yields vastly superior predictions for yaw rate, roll rate, and lateral acceleration compared to traditional quasi-steady curve fits, with results falling well within the standard variability of physical experimental data (SAE, 2017).

Fully Coupled (Two-Way) 6-DOF Simulation: For high-downforce hypercars, where millimeter changes in ride height alter forces by thousands of Newtons, one-way coupling is insufficient because it breaks the feedback loop. Fully coupled, 6-Degree-of-Freedom (6-DOF) simulations create a complete, closed-loop physical system (SAE, 2016). As transient aerodynamics are introduced, the resulting vehicle response (calculated by the MBD handling model) dynamically alters the vehicle's position within the CFD environment (Pacejka, 2012; SAE, 2016).

This requires immensely complex mesh management, typically utilizing the overset grid (chimera) method (SAE, 2016; SAE, 2017). In this approach, a background fluid mesh remains static while a secondary, high-resolution mesh enveloping the vehicle moves freely through the domain according to MBD positional outputs (SAE, 2016; SAE, 2017). In benchmark simulations of the Driver fastback model, differences between quasi-steady, one-way, and two-way coupled methods revealed that transient responses create overshoots in aerodynamic yaw moments, resulting in lateral deviations of up to 8% larger than predicted by steady-state models (SAE, 2016). In the context of high-downforce limit handling, an 8% deviation in yaw moment calculation means the difference between predicting successful, on-the-limit corner negotiation and catastrophic spinout into the barriers.

Coupling Strategy	Data Flow	CFD Mesh Updating Methodology	Computational Cost	Limit Handling Accuracy
Quasi-Steady (Aero Map)	Pre-computed static lookup tables fed to MBD.	None (static configurations pre-meshed).	Low	Poor for rapid transients; good for steady-state corners.
One-Way Transient	CFD forces \rightarrow MBD (time domain).	None (vehicle fixed in CFD space).	Medium	Moderate; ignores aero-structural feedback loop.
Two-Way Fully Coupled	MBD kinematics \leftrightarrow CFD forces (time domain).	Dynamic (Overset Grids / Mesh Morphing).	Extremely High	Excellent; captures full aero-elastic/kinematic feedback.
Table 3: Comparison of aerodynamic-MBD simulation methodologies utilized in the prediction of hypercar limit handling (Carreira, 2020; SAE, 2016; SAE, 2017).				

Mid-Fidelity Approaches and Future Optimization

As the computational expense of two-way CFD-MBD coupling remains extraordinarily high, requiring vast supercomputing clusters, the industry is increasingly leveraging mid-fidelity aerodynamic solvers to expedite the design process. For instance, Vortex Particle Methods (VPM) combined with multi-physics coupling libraries like preCICE are being utilized alongside MBDyn to rapidly evaluate wake modeling and transient aerodynamics without the overhead of full Navier-Stokes volume meshes (PreCICE, 2021).

These mid-fidelity approaches represent a highly effective trade-off, capturing the complex interactional aerodynamics required to evaluate aeroelastic stability, pitch sensitivities, and transient loads (PreCICE, 2021). By establishing these closed-loop design processes, researchers can conduct joint studies on aerodynamic characteristics and handling stability rapidly. In one instance targeting a Formula SAE racing vehicle, the iterative modification of aerodynamic parameters directly influenced by limit handling metrics resulted in an 11.39% increase in downforce and a 2.85% reduction in drag, boosting the overall aerodynamic efficiency (L/D) by 14.70% (MDPI, 2022). More importantly for drivability, the dynamic pitching moment was reduced by 82.34%, massively enhancing the stability of the vehicle platform during transient maneuvers and delaying the onset of limit-handling breakaway (MDPI, 2022).

Methodology

This study employs a mixed-methods research design, integrating primary quantitative and qualitative data gathered through a targeted expert survey with a comprehensive review of recent literature. The primary data collection instrument was a 10-item structured questionnaire

distributed to 33 industry professionals and academic researchers specializing in high-performance vehicle engineering. The respondent pool was carefully segmented to ensure a balanced perspective across multiple technical domains, including experts in Vehicle Dynamics and Multi-Body Dynamics (MBD) (n=10), Aerodynamics and Computational Fluid Dynamics (CFD) (n=8), Tire Modeling (n=5), Chassis and Suspension Design (n=5), and Simulation Software Engineering (n=5).

The survey assessed current industry practices regarding aerodynamic load modeling, simulation software preferences (e.g., ADAMS/Car, SIMPACK, OpenFOAM, STAR-CCM+, IPG CarMaker, and DUST), suspension coupling strategies, and the limitations of empirical tire models during limit handling. To supplement the empirical data, qualitative insights were collected through optional open-ended supplementary comments at the end of the survey. The secondary phase of the methodology involved triangulating these survey findings with peer-reviewed literature on advanced co-simulation architectures, vortex particle methods (VPM), and transient tire dynamics to validate the empirical data and contextualize the current computational bottlenecks.

Data Analysis

The survey responses yielded significant, granular insights into the practical challenges and current bottlenecks of simulating high-downforce limit handling.

Aerodynamic Sensitivities and Suspension Strategies: A significant majority of respondents (72.7%, or 24 out of 33) identified pitch angle (e.g., diving under heavy braking) as the vehicle attitude parameter posing the highest risk for inducing a sudden loss of downforce or aerodynamic instability. Roll angle and yaw/slip angle were identified as the primary risk by only 18.2% and 9.1% of respondents, respectively. To mitigate phenomena such as aerodynamic "porpoising" or high-speed bounce, engineers utilize a variety of mechanical tuning variables: the preferred method is stiffening the heave/third spring (36.4%), followed by increasing the static ride height (30.3%), increasing the bump stop gap to engage later (18.2%), and utilizing softer bump stops with increased damping (12.1%). Consequently, 90.9% of the experts surveyed rated the implementation of decoupled suspension topologies as highly or absolutely essential for maintaining the required aerodynamic platform.

Simulation Architectures: Despite the known limitations in capturing transient events, slightly more than half of the respondents (54.5%) still rely on static quasi-steady aero maps (lookup tables) as their preferred method for integrating aerodynamic data into MBD simulations. Fully coupled (two-way) 6-DOF CFD-MBD co-simulation, one-way transient coupling, and mid-fidelity approaches like Vortex Particle Methods (VPM) were each preferred by only 15.2% of the sample. When asked to evaluate the accuracy of predicting a vehicle's response to transient events (such as a sudden crosswind gust), 42.4% noted that coupled simulations are "Moderately more accurate" (a 5% to 10% deviation difference), and 33.3% found them "Significantly more accurate" (greater than 10% deviation difference).

Computational Bottlenecks: When asked to identify the most significant bottleneck

preventing the widespread adoption of fully coupled (two-way) time-domain CFD-MBD co-simulations, respondents were divided. Extreme computational expense and hardware limits were cited by 36.4% of the experts. This was closely followed by a lack of access to validated, high-fidelity tire models (30.3%), dynamic fluid meshing and overset grid difficulties (18.2%), and software communication stability issues, specifically explicit loose coupling scheme divergence (15.2%).

Tire Load Sensitivity: The inadequacy of legacy tire models at the aerodynamic limit was a point of unanimous agreement. All 33 respondents (100%) indicated that standard empirical tire models (such as the Pacejka Magic Formula) either require heavy parameter modification or must be completely replaced by custom transient models to accurately capture grip degradation during rapid lateral load transfers. Furthermore, 100% of the experts rated tire load sensitivity as having either a "Significant impact" (42.4%) or fundamentally dictating the accuracy of handling models (57.6%) at the aerodynamic limit.

Interpretation

The empirical data strongly suggests that managing the aero-mechanical platform is the most critical factor in hypercar limit handling, directly mirroring recent aerodynamic literature. The survey's finding that pitch angle is the most dangerous parameter aligns with fluid dynamics research demonstrating that race cars and high-downforce vehicles are exceptionally sensitive to ride-height variations. As a high-downforce vehicle pitches forward under heavy braking, the airflow entering the underbody changes rapidly; the distance between the front wing and the ground decreases, altering the front wing tip vortices and ground effect structures. If the ride height drops below a critical threshold, it can trigger vortex breakdown inside the diffuser, leading to a sudden, catastrophic loss of rear downforce and severe snap oversteer. This physical reality explains why nearly all surveyed experts mandate the use of decoupled suspension systems (such as heave springs or third dampers) to strictly control pitch and heave motions independently of vehicle roll.

Furthermore, a significant gap exists between the theoretical ideal of vehicle simulation and actual industry practice. While respondents widely acknowledged that coupled CFD-MBD simulations are significantly more accurate for predicting responses to transient events, extreme computational costs and dynamic meshing difficulties keep the industry largely reliant on quasi-steady aero maps. However, qualitative feedback from the survey noted an emerging solution: mid-fidelity tools like DUST, which utilize a Vortex Particle Method (VPM). VPMs offer a highly effective middle-ground fidelity, successfully bridging the gap between static lookup tables and overly expensive Unsteady Reynolds-Averaged Navier-Stokes (URANS) models. Because VPM is a Lagrangian grid-free method that describes the evolution of the wake through the rotational component of velocity, it sidesteps the severe overset mesh deformation issues that plague traditional CFD solvers during large suspension displacements.

Finally, the unanimous consensus regarding the inadequacy of empirical tire models highlights a critical area for future engineering development. In high-downforce applications,

aerodynamic forces are utilized to drastically increase the vertical load on the tires, thereby increasing the maximum possible lateral friction for cornering. However, as vertical load increases to extreme levels, the tire's coefficient of friction degrades non-linearly—a phenomenon known as tire load sensitivity. Empirical models like the Pacejka Magic Formula, which rely on single-point contact assumptions, struggle to mathematically capture this highly transient grip degradation. To accurately sign-off on limit-handling maneuvers, engineers are increasingly turning to advanced physical, structure-based tire models, such as FTire, integrated into multibody solvers like SIMPACK. These structure-based models account for the high-frequency dynamics (up to 250\text{ Hz}), severe load sensitivities, and complex thermal and thermodynamic effects that define the tire-road interaction of a modern hypercar at the absolute limit of adhesion.

Recommendations

Based on the integration of industry survey data and contemporary literature, the following recommendations are proposed for high-performance automotive engineering teams: *Transition to Mid-Fidelity Aerodynamic Solvers:* Teams should bridge the gap between static quasi-steady aero maps and computationally prohibitive URANS models by adopting mid-fidelity solvers based on Vortex Particle Methods (VPM), such as DUST. This approach provides a necessary increase in transient accuracy without the severe meshing penalties of fully coupled two-way CFD-MBD co-simulations.

Adopt Structure-Based Tire Models for High-Downforce Validation: The unanimous agreement on the limitations of empirical tire models dictates that Pacejka-based models should be restricted to low-speed or steady-state analyses. For limit-handling and high-speed transient maneuvers, engineers must transition to structure-based tire models (e.g., FTire) that can accurately compute non-linear tire load sensitivities and high-frequency grip degradation.

Mandate Decoupled Suspension Topologies: To safely mitigate extreme pitch sensitivity and the associated risk of diffuser stall, the use of decoupled suspension architectures (featuring independent heave springs and strictly tuned nonlinear bump stops) must be considered a foundational design requirement rather than an optional tuning tool

Limitations

While this study offers valuable insights, several limitations must be acknowledged. First, the sample size of the expert survey (n=33) is relatively small, albeit highly specialized. While sufficient for identifying consensus trends within a niche field, a larger cohort could provide more granular statistical insights. Secondly, the highly confidential nature of motorsport and hypercar development means that some of the most advanced, proprietary co-simulation workflows and custom-developed transient tire models are restricted from public academic review. Consequently, this research relies heavily on commercially available software benchmarks (e.g., SIMPACK, ADAMS, STAR-CCM+) and may not fully capture proprietary in-house solutions utilized by top-tier Formula 1 teams or elite hypercar manufacturers.

Conclusion

Predicting the limit handling of high-downforce hypercars represents a paradigm shift in automotive engineering, where mechanical multi-body dynamics and aerodynamics can no longer be treated as isolated disciplines. As aerodynamic loads push vertical tire forces to unprecedented extremes, traditional simulation tools, specifically static aero maps and empirical Pacejka tire models, fail to capture the highly transient phenomena that dictate a vehicle's stability at the edge of adhesion. This study confirms that pitch-induced aerodynamic instability and severe tire load sensitivity are the primary limiting factors in modern hypercar performance. While computational expense remains a significant bottleneck for fully coupled co-simulations, the strategic integration of mid-fidelity Vortex Particle Methods, advanced structural tire models, and strictly decoupled mechanical suspension topologies provides a robust framework for safely engineering the next generation of extreme-performance vehicles.

Works Cited

Bauchau, O. A., Wu, G., Betsch, P., & Sonneville, V. (n.d.). Tip twist vs. loading angle for three loading conditions. ResearchGate.

cousin Scientific Software. (2021). FTire: The multi purpose virtual tire model [Brochure]. cosin.eu.

Dassault Systèmes. (n.d.). Full Vehicle Analysis with Simpack Automotive. SIMULIA.

Katz, J. (2006). Aerodynamics of race cars. Annual Review of Fluid Mechanics, 38, 27-63.

Ma, X., Li, J., Zhang, K., Zou, Y., & Massaro, M. (2026). Effects of wheel-ground conditions on racing car aerodynamics under ride-height-related attitude variations. Applied Sciences, 16(2), 874.

Montagnani, D., & Savino, A. (2019). DUST: A novel flexible mid-fidelity aerodynamic solver for rotary-wing vehicles. Politecnico di Milano.

Savino, A., Cocco, A., & Zanotti, A. (2022). Numerical investigation of wing-propeller aerodynamic interaction through a vortex particle-based aerodynamic solver. Proceedings of the 48th European Rotorcraft Forum.

Wang, Y., Li, X., Qiao, G., et al. (n.d.). Co-simulation and explicit loose coupling in vehicle aerodynamics and MBD. ResearchGate.

Zhang, X., & Lygeros, J. (n.d.). Tire and vehicle dynamics: Co-simulation architectures. ResearchGate.

Comparing Narrative Recall After AI Summarization and Full-Text Reading

By Rohan Murthy

Abstract

This study examines differences in students' narrative recall performance following exposure to literature with an AI-generated summary or the original full-text reading. This research study intends to address the topic by assessing how well middle school students (N = 96) recall a family friendly science fiction story after reading either the original text or an AI-generated summary. The use of AI is becoming even more widespread in the educational sphere. Spurred by the College Board's lenient stance on generative AI, I conducted this study to see if the time saved by reading AI summaries compromises information retention.

Understanding these drawbacks is critical, as memory and recall are essential for student academic success. To assess this, students were first tasked with reading the assigned text. Two days later the students were given a short quiz with a varying difficulty of questions on the story. As sleep is a recognized factor essential in the consolidation of memory, I analyzed it as a potential confounding variable and accounted for it by collecting information on the students' sleep duration and quality the previous night. By addressing this topic, I intend to analyze how AI summary tools could impact learning and memory in an educational setting. Following the study, I saw a 6.9% difference in the scores, with the group who read the AI summary scoring on average better on this assessment than the group who read the original text. One possible explanation is that well-written summaries may make key inferences more explicit, making them easier to retrieve during later recall. Then, rather than the student having to do the cognitive processing themselves, they can just recall those surfaced inferences that have been made for them two days later.

KeyWords Artificial Intelligence, Psychology, Memory Retention, Sleep, Summarization, sleep duration, middle school students, narrative recall, generative ai, sleep quality, chatgpt

Introduction

Artificial intelligence (AI) has quickly become an integral part of our daily life, including the educational sphere. AI tools are being utilized by students all over the world to aid in summarizing reading, explaining concepts, and even fully completing homework. A recent survey uncovered 84% of students have utilized a generative AI tool (College Board, 2025), with the most popular choice being ChatGPT, to aid them in their schoolwork. This underscores the significant role AI has already taken on in the educational system. Although AI tools can improve accessibility and provide personalized support, their impact in deeper learning and memory recall still remains unclear. "High school students, parents, and educators believe there are benefits to using AI, but they also have concerns" (College Board, 2025). Research into AI learning tools has shown that they may have disadvantages. As Torney notes, "These large language models aren't necessarily optimized for the task of learning, which is slower [and]

takes effort. In general, [we] have seen a little bit of a slower uptake among officially sanctioned school uses of AI with kids" (Vilcarno, 2025).

To understand the impact of AI tools used for summarization on the learning process, it is important to understand how memory is affected. The process of learning is heavily dependent on a person's memory, especially their long-term memory. Cognitive psychologists have established that the human memory is composed of working, declarative, and non-declarative memories (Sridhar 2023). Working memory refers to the short-term memory system involved in temporarily holding and manipulating information for the performance of activities such as reading or problem-solving (Baddeley 1992). Information processed by the working memory can be stored in long-term memory through learning and repetition. Short-term episodic memory, which is a temporary form of episodic memory, can affect the ability to recall recently experienced events, such as reading a story or a summary (Gavett 2011). All these types of memory, though, are thoroughly influenced by sleep (Yale 2022).

By accounting for variations in sleep duration and quality, I can further understand what part of the correlation is due to a difference in treatments, or confounding variables, like sleep. Sleep aids memory consolidation, and additionally sleep deprivation has a detrimental effect on memory (Gais et al., 2006).

Changes in the way information is processed can affect a person's ability to retain information. When students use summaries created by AI to attempt to learn the information, instead of reading the original texts, the way the information is processed in memory is affected. It has been observed that AI summarization tools reduce engagement with source material, encouraging surface-level processing, which correlates with poorer comprehension depth (Putri 2025). This leads us to wonder if a person's ability to retain information after reading an AI summary of a text is equivalent to a person who reads the original full text.

This study examines the relationship between text presentation format and reading comprehension by comparing how well middle school students recall a 1,038-word retro-futuristic children's story after reading either the original text or an AI-generated summary. Additionally, the analysis accounts for participants' sleep duration, recognizing its established role in memory consolidation. I hypothesize that students who read the original story will, on average, achieve higher scores on the recall quiz than those who read the AI summary, as the full text may engage deeper cognitive processing.

Literature Review

The process of learning and memory recall is influenced by a range of external factors, including sleep quality and the cognitive demands placed on learners during study. To understand the implications of the present study, it is important to review the existing research on how sleep affects memory consolidation, and how the use of AI tools affects the depth of learning and information recall.

Conte et al. and Newbury et al. investigated the relationship between sleep and memory through differing methods, focusing on the preservation and encoding of information. Conte et

al. targeted the consolidation of complex narrative texts by having university students learn prose passages and testing them twice: once immediately and again after a nine-hour "recall interval" where participants were either asleep or awake. In contrast, Newbury et al. provided a broader quantitative scope by conducting two meta-analyses of studies published over fifty years, examining the impact of total acute sleep deprivation both before and after the learning process. While Conte et al. focused on the specific decay of narrative memory, Newbury et al. measured effect sizes across various memory types, finding that deprivation before learning had a more severe impact (Hedges' $g = 0.621$) than deprivation occurring after learning (Hedges' $g = 0.277$). Despite these different approaches, both studies lead to the unified conclusion that sleep is an active and essential component of the learning cycle rather than a passive state of rest. The findings collectively suggest that sleep serves two vital functions: it prepares the brain to encode new information and subsequently stabilizes that information by counteracting the natural "decay" or loss of knowledge that occurs during wakefulness. For the purposes of the present study, these conclusions underscore that sleep quality must be treated as a critical covariate; a participant's inability to recall details from a story may be a direct result of poor sleep hindering memory consolidation, independent of whether they read the original text or an AI summary (Conte et al., 2025) (Newbury et al., 2021).

Crowley et al. researched further into the niche of the far more common scenario of students getting partial sleep restriction rather than going entirely without sleep. Their study sought to determine whether sleeping for fewer hours than recommended carries meaningful consequences for long-term information recall. The study followed 56 adolescent students who were restricted to either 5 hours or 9 hours of sleep per night over a period of five nights. Students learned factual material each day and were tested on recall at multiple intervals, including one week later. Restricting sleep to between 3 and 6.5 hours, compared to normal sleep of 7 to 11 hours, negatively affected memory formation with a small but statistically significant effect size (Hedges' $g = 0.29$, 95% CI [0.13, 0.44]), and there was no statistically significant difference found when these effect sizes were analyzed and compared with those from total sleep deprivation studies, meaning any difference was likely due to random chance, and suggesting that missing some sleep has similar consequences for memory as not sleeping at all. These findings are important to understand for our study, as the target demographic (middle school students) are much more likely to experience moderate sleep length (Crowley et al., 2024).

To bridge the connection between physiological factors like sleep and the cognitive impact of technology, it is essential to consider how external tools also shape the learning process. While sleep provides the biological foundation for memory consolidation, the way students interact with information—increasingly through the use of artificial intelligence—introduces new variables into how that information is initially encoded. As students transition from traditional reading to AI-assisted methods, the cognitive effort involves changes, leading to a phenomenon known as "cognitive offloading". This shift from internal biological processes, to external digital assistance, necessitates a closer look at how AI integration specifically influences cognitive engagement and long-term recall.

A study performed by Michael Gerlich at SBS Swiss Business School identified a growing concern that, while the productivity benefits of AI tools were widely discussed, their effects on deeper cognitive processes, particularly critical thinking, remained poorly understood. Gerlich set out to investigate whether frequent AI tool usage was associated with measurable declines in critical thinking ability, and whether it was specifically as a result of cognitive offloading — when individuals rely on the tools to reduce mental effort. Quantitative data collection involved measuring AI tool usage, cognitive offloading tendencies, and critical thinking skills with a 23-item questionnaire. 50 of the participants underwent semi-structured interviews, which has solid thematic analysis as qualitative data. The results demonstrated that cognitive offloading was strongly correlated with AI tool usage ($r = +0.72$) and inversely related to critical thinking ($r = -0.75$). Younger participants aged 17-25 showed higher dependence on AI tools and scored lower on critical thinking assessments compared to older age groups. This study is relevant because it suggests that students who habitually rely on AI tools may already engage in shallower cognitive processing and critical thinking even when they do attempt to read and absorb material (Gerlich, 2025).

Qirui Ju aimed to quantify the impact of generative AI tools on learning in academic reading materials. The central problem was that while AI tools were increasingly being used by students to assist with reading and writing tasks, it remained unclear how much that reliance was actually costing them in terms of comprehension accuracy. College-educated participants engaged in three cycles of reading and writing tasks, after which they answered comprehension questions about the material. The design allowed for comparison across three conditions: no AI assistance, AI-assisted reading, and complete reliance on AI for writing tasks. After adjusting for background knowledge and demographic factors, complete reliance on AI for writing tasks led to a 25.1% reduction in accuracy. In contrast, AI-assisted reading resulted in a 12% decline. Further analysis revealed that individuals with a robust background in the reading topic and superior reading and writing skills benefited the most from AI assistance. The significance of these findings for the present study is a warning that even when students use AI only to assist with reading there is a measurable reduction in how accurately they recall and apply the information afterward (Faber 2023). Although this is similar to our own study, ours has several key differences. The demographics vary, with ours being tailored towards middle school students, and also has 2 categories / treatments plants versus the 3 in Ju's study. Along with that, our study takes into account sleep as a variable affecting the outcome, and asks the questions two days later instead of during the same day.

Materials

I used ChatGPT Pro (Model GPT-5.3) to summarize the text and I found the full text of *"The Fun They Had"* online. I printed both full text and generated summaries on paper. Finally, I wrote down quiz questions and printed them out along with sleep surveys.

Methodology

To compare information recall between students reading full literature versus AI summarized text, this study utilized a between-subjects design where middle school students were allocated to one of two conditions: reading the full text or reading an AI-generated summary. I selected Isaac Asimov's short story *"The Fun They Had,"* as the stimulus selected for its brief length of two pages.The story also has thematic elements suitable for generating a variety of factual and analytic questions . ChatGPT Pro (Model GPT-5.3) on March 3, 2026 generated a summary based on a standardized prompt I provided to match the two-page constraint of the original text. This fixed length controlled for variations in compression depth and information density, ensuring a uniform summarization treatment rather than unconstrained model generation.

The experiment was conducted over two sessions, spaced two days apart. During the initial session, participants were assigned to Group A (original text) or Group B (AI summary) via coin flip. Approximately 100 students across two to three class periods were allocated to each condition. Following the oral reading of the informed consent disclosure, students were given the option to opt out before stimuli were distributed face-down. Participants were given exactly 10 minutes to read their assigned text, after which materials were collected with instructions prohibiting peer discussion. Two days later, in the second session, students were once again read the informed consent disclosure before being given 10 minutes to complete a demographic survey and comprehension quiz with questions from multiple cognitive levels. The survey gathered data regarding sex assigned at birth and self-reported sleep duration from the previous night, recognizing sleep as a critical variable in memory consolidation.

This structure, however, introduced several methodological limitations and threats to internal and ecological validity that must be acknowledged. Because the formal consent process was conducted out loud immediately prior to testing in a structured environment, participants were explicitly aware of their status in a research study. This introduces the potential for the Hawthorne effect, where behavioral reactivity and heightened situational motivation, rather than the text format alone, may heavily influence a student's willingness to recall information. Furthermore, administering the experiment within an established class-period framework introduces a clustering effect. This presents an internal validity threat, as the data may be confounded by non-random environmental variables unique to each period, such as time-of-day fatigue and localized peer-group dynamics.

The physical constraints of the text and time limits also present distinct challenges. Utilizing a brief, two-page text reduces the ecological validity of the study, as typical educational assessments generally require navigating denser, more complex texts. This structural constraint that comes with shorter text may inadvertently grant an artificial advantage to the AI condition, as the summary efficiently strips "fluff" and compresses the material into highly recallable units. Finally, while the 10-minute reading window was strictly standardized, equal time does not guarantee equal cognitive pressure. The study could not control for individual variances in baseline reading speed, active engagement, or classroom distractions like ambient noise,

meaning the actual cognitive load experienced during the task likely varied significantly across participants.

To assess recall across different cognitive depths, the 14-question quiz was tiered into three distinct levels of cognitive processing. All questions were curated to test only information present in both the AI summary and the original story to ensure equitable access to the material. The first seven questions targeted lower-order cognitive processes, specifically literal recall and recognition memory, requiring students to retrieve simple facts from semantic memory (e.g., "What was the name of the main character?") via surface-level processing and direct recall pathways. The subsequent five questions assessed intermediate-level comprehension, targeting inferential processing and relational understanding. These items required students to integrate multiple pieces of information, engaging working memory and semantic integration to generate connections beyond explicitly stated details (Alptekin, 2011). The final two questions targeted higher-order cognitive processes, specifically deep semantic processing, abstract reasoning, and conceptual integration. These questions required students to engage in elaborative rehearsal and activate long-term semantic networks to evaluate overarching themes (e.g., "What idea about learning does the story suggest?"), reflecting deep processing strongly associated with knowledge transfer (Craik, 2002).

Following data collection, quiz performance was graded and the resulting dataset was screened for anomalies. Data entries with missing points or those completed outside the standardized time window were excluded as outliers before executing calculations. Data analysis involved graphing the respective datasets against one another to visually inspect for emergent trends and correlations. Descriptive statistics were calculated for each treatment group to compare central tendencies and variances. Finally, a two-sample t-test for a difference in population means was conducted to determine if there was a statistically significant difference in information recall between the group exposed to the original text and the group exposed to the AI-generated summary.

After cleaning the dataset, I was ready to analyze the results. The next section conducts and interprets the final analysis, with comparison of the total quiz scores between the reading conditions, and then a comparison after breaking the quiz down into the sections which represent different cognitive levels. Finally, I examined the role of sleep, in terms of raw duration and a composite score compiled from self-reported sleep quality and duration. Together, these analyses test whether the choice between reading the original story and reading an AI-generated summary has a measurable effect on information recall two days later, and whether any observed difference can be explained away by differences in sleep.

Results and Discussion

Of the 121 responses collected, 25 were removed during cleaning. There were 19 submissions that fell outside the time window the experiment was held in, four with missing answers or sleep data, and two statistical outliers by the Two Standard Deviation Rule for

outliers. The final analysis contained N = 96 students: 38 in group A (Normal Story) and 58 in Group B. The imbalance is explained by the random assignment at the beginning of the study.

The statistical analysis revealed a significant difference in performance between the two text presentation conditions, shifting in the opposite direction of the original working hypothesis. Group A (original story) achieved a mean total quiz score of 10.21 out of 14 (SD = 1.85), which corresponds to an accuracy rate of 72.9%. In comparison, Group B (AI summary) demonstrated a higher mean total quiz score of 11.17 out of 14 (SD = 1.57), representing an accuracy rate of 79.8% (see Figure 1.). A two-sample t-test for a difference in population means was conducted on the total scores to evaluate these differences. The observed mean difference between the group A and B was 0.96 points on the 14-point scale (95% CI = (0.24, 1.69), t = 2.65, df= ~70.1, p = 0.01), yielding a moderate effect size (Cohen's d = 0.57). Based on these inferential statistics, the null hypothesis of equal mean recall across conditions is rejected at the $\alpha = 0.05$ significance level. Students assigned to the AI-summary condition scored approximately 6.9 percentage points higher than those in the original-text condition. However, this group difference implies better delayed short-term recall for a brief, two-page text rather than serving as an indicator of robust conceptual learning. However, this statistical inference strictly evaluates the variance between sample means; it does not isolate a direct causal treatment effect, as external environmental or behavioral confounding variables may account for the observed difference.

Group	Number (N)	Mean (out of 14)	SD	Mean percentage	Mean sleep hours
A (Normal Story)	38	10.21	1.85	72.9%	7.16
B(AI Summary)	58	11.17	1.57	79.8%	7.69

Figure 1. Descriptive Statistics by treatment

Because the quiz was designed to probe three distinct cognitive levels — lower-order literal recall (Q1-Q7), intermediate-order inferential comprehension (Q8-Q12), and higher-order conceptual integration (Q13-Q14) — the overall t-test was complemented by two-sample z-tests for the difference in proportions at each level (See Figure 2.).

Cognitive Lvl	Group A percentage	Group B percentage	Diff (A-B)	z	p-value
Lower (Q1-Q7)	62%	64%	-2%	-0.53	0.597
Intermediate (Q8-Q12)	80.5%	93.8%	-13.3%	-4.46	< 0.001
Higher (Q13-Q14)	92.1%	100%	-7.9%	-3.07	0.002
Overall (Q1-Q14)	72.9%	79.8%	-6.9%	-2.93	0.003

Figure 2. Accuracy and two sample z tests by cognitive level

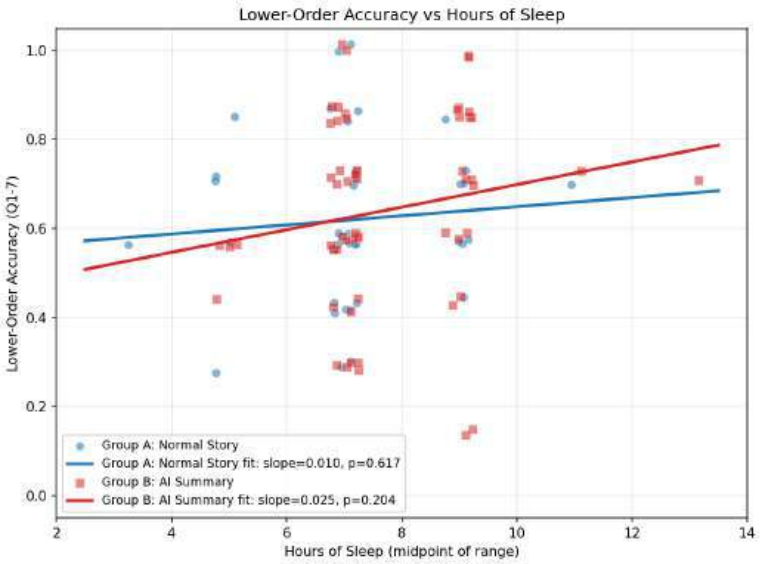


Figure 3. Lower-order questions score (percentage) vs. hours of sleep, by treatment.

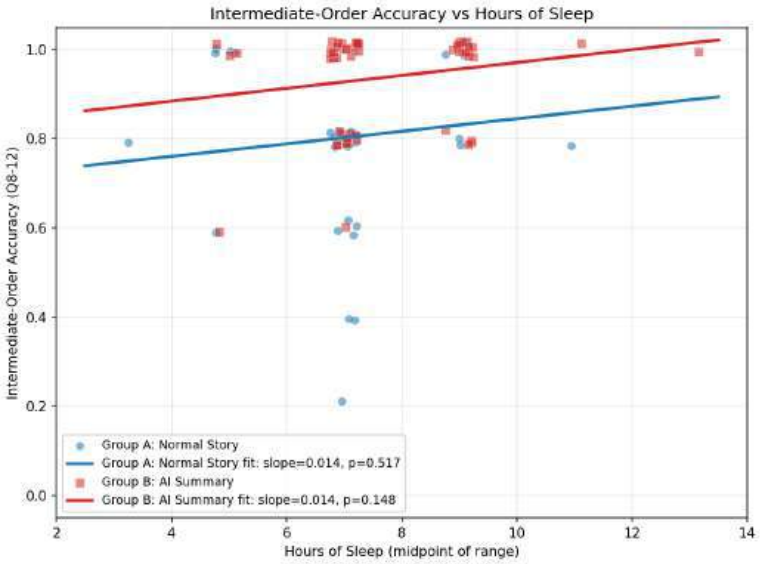


Figure 4. Intermediate-order questions score (percentage) vs. hours of sleep, by treatment.

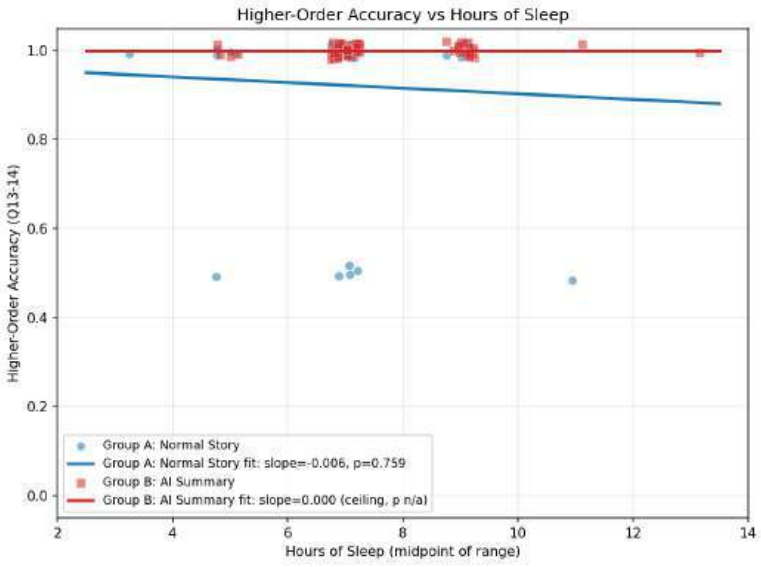


Figure 5. Higher-order questions score (percentage) vs. hours of sleep, by treatment.

The lower-order recall difference (See Figure 3.) is small (62.0% vs. 64.0%, $z = -0.53$, $p = 0.597$) and not statistically significant. Surface facts such as character names, the year, and the type of teacher appear to be encoded almost equally well from the original story and the AI summary. The remaining differences, however, are substantial. At the intermediate level (See Figure 4.), Group B answered 93.8% of items correctly compared to 80.5% in Group A, which is a 13.3 percentage-point gap ($z = -4.46$, $p < 0.001$). At the higher-order level (See Figure 5.), Group B achieved a perfect score (100%), compared with 92.1% for Group A ($z = -3.07$, $p = 0.002$). In either reading, the AI-summary group answered every inferential thematic question correctly. Across all 14 items, the overall accuracy gap was 6.9 percentage points (79.8% vs. 72.9%; $z = -2.93$, $p = 0.003$).

These cognitive-level results explain why the overall t-test detected a statistically significant difference despite the small 0.960-point mean difference: the observed advantage for the AI-summary group is concentrated in the intermediate and higher order questions, not in lower order ones. This concentrated advantage on higher-order performance warrants closer examination, particularly given that the AI-summary group achieved 100% accuracy on the higher-order questions. Rather than indicating superior conceptual learning or advanced critical analysis, this perfect score suggests a potential aspect of the summarization itself. In the original narrative text, overarching themes and abstract ideas are embedded implicitly within the text, requiring readers to execute cognitive processing to infer them. Conversely, the AI-generated summary likely converted these implicit themes into explicit, obvious statements. Consequently, for Group B, the final two quiz questions ceased to function as assessments of abstract reasoning or deep cognitive processing; instead, they were reduced to simple recognition and literal recall tasks. This indicates that while students in the AI-summary condition demonstrated near-perfect retrieval of thematic details, it did so by giving away the analytical answers directly, thereby

bypassing the cognitive friction essential for genuine knowledge transfer. Notably, this result runs counter to Ju's finding that AI-assisted reading reduced comprehension accuracy by 12% relative to unassisted reading (Faber 2023).

The literature reviewed above (Conte et al., 2025; Crowley et al., 2024; Newbury et al., 2021) makes it clear that sleep is critical for any memory mechanisms. The post-survey collected both sleep duration (in two-hour bins from 0-2 to 12+) and a self-reported sleep-quality rating on a 1-10 Likert scale. Sleep duration was coded as the midpoint of the reported range (0-2 = 1, 2-4 = 3, …, 12+ = 13). To examine how sleep relates to recall within each reading condition, total quiz score was first regressed on hours of sleep (See Figure 6. and Figure 7.), separately for each treatment, using ordinary least squares.

Group	Number (N)	Slope (points / hour)	Intercept	p-value	R^2
A (Normal Story)	38	0.129	9.29	0.531	0.011
B (AI Summary)	58	0.25	9.25	0.086	0.052

Figure 6. Within-group linear regression: total score on hours of sleep

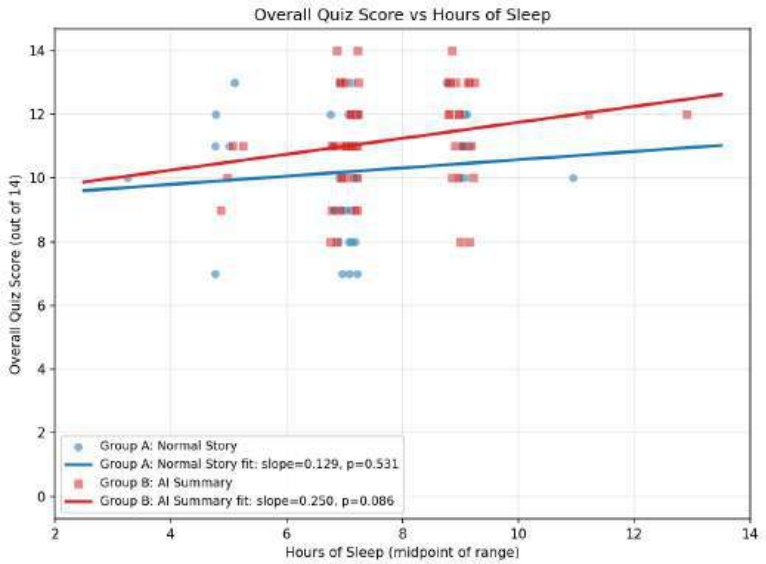


Figure 7. Overall quiz score (out of 14) vs hours of sleep, by treatment.

Both slopes are positive — more sleep is weakly associated with higher quiz performance — but neither slope reaches statistical significance at α = 0.05. The Group B (AI Summary) trend is closer to the threshold (b = 0.250 points per additional hour, p = 0.086, R^2 = 0.052) than the Group A trend (b = 0.129, p = 0.531, R^2 = 0.011). The small R^2 values mean that within either group, sleep duration accounts for 5.2 and 1.1 percent respectively of the variance that is observed in total score; the dominant variance is between-student. This is qualitatively consistent with Crowley et al.'s small-but-real effect size for partial sleep restriction.

Sleep duration is only one component of sleep. To additionally check whether self-reported sleep quality can predict recall, the same simple regression was rerun with sleep quality (1-10) as the predictor (See Figure 8.), and also run with a composite sleep score (See Figure 9.) that was constructed by taking the mean of the z-standardized values of duration and quality. Sleep duration and quality were only weakly correlated in this sample (r = 0.26), so the composite reflects two distinct facets of sleep rather than a single redundant signal. Neither sleep quality alone nor the composite score was a significant within group predictor of total quiz score (sleep quality: Group A slope = -0.04, p = 0.798; Group B slope = +0.01, p = 0.900; composite: Group A slope = +0.10, p = 0.806; Group B slope = +0.33, p = 0.238). In other words, the weak sleep-recall trend observed in this study is carried by sleep duration, not by self-reported sleep quality, and combining the two into a composite does not produce a stronger signal.

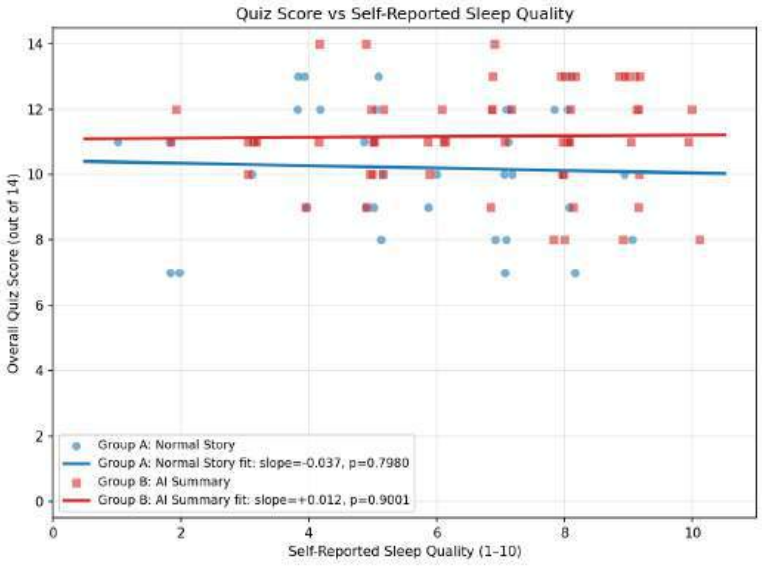


Figure 8. Overall quiz score (out of 14) vs self-reported sleep quality, by treatment.

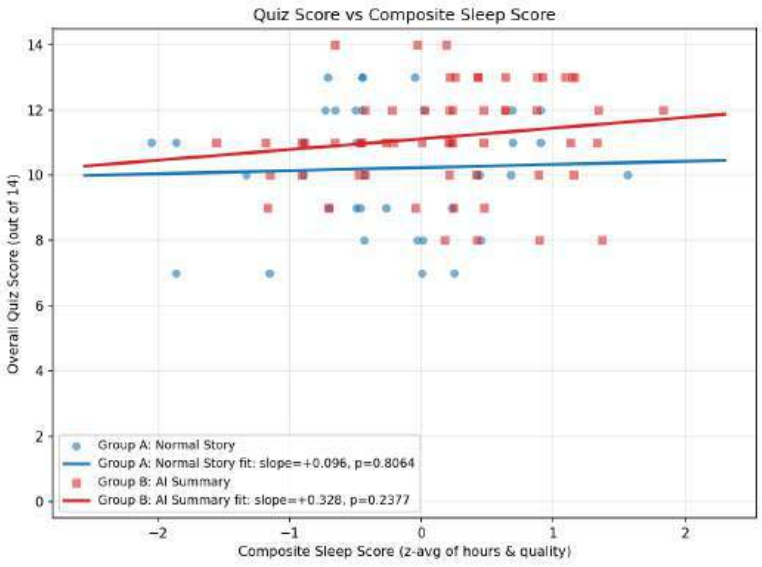


Figure 9. Overall quiz score (out of 14) vs composite sleep score (avg of z-standardized values of sleep quality and duration), by treatment.

To test whether the treatment effect on quiz score survives once sleep is controlled for, three pooled multiple-regression models were estimated on the full N = 96 sample, using a treatment indicator (1 for AI Summary, 0 for Normal Story). The three models address their own related but distinct questions. Model A (See Figure 10.) includes sleep duration and self-reported sleep quality as separate predictors alongside the treatment indicator, allowing each component of sleep to compete on its own merits and, fundamentally, asking whether the AI-summary advantage survives once both sleep variables are statistically held constant in order to test if some of the differences were due to differences in sleep. Model B (See Figure 11.) replaces the two separate sleep predictors with a composite sleep score— calculated by taking the average of the standardized z-score of both sleep variable— in order to test whether combining duration and quality into a single z-averaged value more efficiently captures how 'well rested' an individual was than treating both variables separately. Model C (See Figure 12.) extends Model A with two interaction terms — sleep hours x Group_B and sleep quality x Group_B — and is the formal test of whether the relationship between sleep and recall differs by reading condition. A significant interaction would indicate that, for example, sleep matters more for students who read the original story than for students who read the AI summary. By contrast, a non-significant interaction implies that the per-group slope differences in Table 3 are within sampling noise. Together, the three models let us separate three questions that are easy to conflate: whether sleep helps explain recall at all (Models A and B), whether a single composite measure of sleep is preferable to its two components (Model A versus Model B), and whether the sleep effect itself depends on which version of the text the student read (Model C).

Model	Estimate	SE	T	P-value
Model A				
Intercept	+8.911	0.9308	+9.57	< 0.001
Sleep Hours	+0.2132	0.1218	+1.75	0.083
Sleep Quality	-0.04	0.0827	-0.48	0.63
Group B	+0.8933	0.3625	+2.46	0.016

Figure 10. Pooled multiple-regression results for Model A (outcome is total quiz score)

Model	Estimate	SE	T	P-value
Model B				
Intercept	+10.2711	0.2795	+36.75	< 0.001
Composite Sleep Score	+0.231	0.2258	+1.02	0.309
Group B	+0.8616	0.3648	+2.36	0.02

Figure 11. Pooled multiple-regression results for Model B (outcome is total quiz score)

Model	Estimate	SE	T	P-value
Model C				
Intercept	+9.4781	1.411	+6.72	< 0.001
Sleep Hours	+0.155	0.1927	+0.8	0.423
Sleep Quality	-0.0666	0.1352	-0.49	0.623
Group B	-0.1312	1.9277	-0.07	0.946
Hours x Group B	+0.1015	0.2504	+0.41	0.686
Quality x Group B	+0.0449	0.1719	+0.26	0.794

Figure 12. Pooled multiple-regression results for Model C (outcome is total quiz score)

Across all three pooled models the treatment indicator is the only statistically significant predictor of total quiz score. In Model A, students in the AI-summary condition score 0.89 points higher on the 14 - point quiz than students in the original-story condition after controlling for

both sleep duration and sleep quality (β = +0.89, SE = 0.36, t = +2.46, p = 0.016). Model B, which replaces the two sleep predictors with the composite sleep score, gives essentially the same treatment estimate (β = +0.86, p = 0.020) while the composite slope is not significant (p = 0.309). Model C adds treatment x sleep interactions to test whether the sleep-score relationship differs by reading condition; both interaction coefficients are not significant (p = 0.686 and p = 0.794), indicating that the apparent slope difference between groups in Table 3 is well within sampling noise. The headline implication is that the AI-summary advantage on recall is not explained by sleep differences between the two groups, and that the sleep factor, while of theoretically important and the correct sign for duration, contributes little explanatory power in this sample beyond the treatment indicator.

Conclusion

This study set out to answer a single question: when middle school students read an AI-generated summary of a short story in place of the original text, is their ability to retain information about that story two days later equivalent to the ability of students who read the full text? Across a final sample of 96 students drawn from five class periods (38 in the original-story condition, 58 in the AI-summary condition), the answer is there is an effect on recall, but the direction of the difference is the opposite of what the working hypothesis predicted. Students who read the AI summary scored on average 0.96 points higher on the 14 - question quiz than students who read the original story (79.8% vs. 72.9%; Welch's t = 2.65, p = 0.010, Cohen's d = 0.57). Separating the quiz questions by cognitive level showed that the gap is concentrated at the intermediate inferential level (Q8-Q12, +13.3 percentage points, p < 0.001) and at the higher-order thematic level (Q13–Q14, 100% vs. 92.1%, p = 0.002), while literal recall questions (Q1–Q7) were answered equally well in both groups (p = 0.597). In other words, an AI-generated summary appears to make the story's inferential and thematic content more retrievable two days later for short stories like the one used in the study, without sacrificing recall of surface facts.

Sleep, the confounding variable analyzed, played a much smaller role than the literature might lead one to expect. Within each reading condition, the total quiz score showed a statistically insignificant positive relationship with hours of sleep (Group A slope = 0.13 per hour, p = 0.531; Group B slope = 0.25, p = 0.086). Self-reported sleep quality, on a 1–10 Likert scale, showed essentially no relationship in either group (both p > 0.79). A composite sleep score constructed by averaging z-standardized duration and quality did not improve on either predictor alone. When sleep duration, sleep quality, and the treatment indicator were entered jointly in a pooled multiple-regression model, the treatment indicator was the only significant predictor of total score (b = +0.89 points, p = 0.016). Sleep duration retained the expected positive sign but was only marginal (p = 0.083). Interaction terms between treatment and the sleep predictors were far from significant, indicating that the sleep-score relationship is statistically indistinguishable between the two reading conditions. The observed AI-summary advantage on recall two days later is therefore not explained by sleep differences between the groups.

These findings should be interpreted alongside several limitations. First, although the student-level randomization on day one followed the protocol, treatment was assigned at the class-period level for the analysis, which confounds reading condition with time of day, and difference in peers; in principle, some of the 0.96-point advantage could be a period effect rather than a pure treatment effect. Second, sleep was reported in two-hour bins and via a single Likert item, both of which add measurement error and reduce power to detect what the meta-analyses of Newbury et al. and Crowley et al. would predict to be a real but small effect. Third, the study used a single short story, “The Fun They Had”, so it is unclear whether the AI-summary advantage generalizes to longer or more conceptually dense texts, where summarization is known to remove more details and nuance. Additionally, due to the random sampling, the AI-summary group had 20 more people in it, meaning one bad score for that group would affect the mean score less than a bad score for the original-text group. This could affect the difference observed when taking into account the inherent knowledge gaps that are present among middle school students. As mentioned, some of the difference could be a period effect, meaning the ‘smarter’ children happened to be concentrated in one group. Finally, the outcome is a short, factual, multiple-choice quiz administered two days after exposure; it measures short-term recall, not deeper or longer-term recall and transfer.

The educational implications of these findings present a more nuanced perspective than either the popular framing of AI as a uniform threat to learning or its characterization as an unconditionally beneficial study aid. In the specific context of a brief narrative text evaluated via a delayed multiple-choice assessment, reading an AI-generated summary resulted in measurably higher recall of inferential and thematic content than reading the original story, with no detectable decrement to lower-order cognitive tasks. While these findings directly contradict the original working hypothesis, they can be interpreted through established cognitive frameworks. As conceptualized by Sweller’s Cognitive Load Theory, a highly structured AI summary may significantly reduce intrinsic cognitive load within a strict, uninterrupted 10-minute reading session. By performing analysis on the text and surfacing core narrative connections, the summary minimizes the working memory capacity required. Furthermore, the condensed presentation likely enhances gist-based encoding, which allows students to form a robust, high-level mental representation of the text's themes. By condensing the main ideas, the AI summary makes key themes readily accessible for retrieval two days later, bypassing the deeper, unguided cognitive processing required by the full text.

Consequently, this study provides thematically clear evidence of AI-generated summaries being associated with higher short-term recall performance in this sample under constrained conditions. However, these immediate performance gains do not account for the broader developmental ramifications of long-term reliance on automated tools. Offloading structural and inferential analysis to an AI model risks bypassing the essential cognitive friction necessary for developing independent reading comprehension skills. As demonstrated in prior research by Gerlich and Jackson, habitual AI reliance strongly correlates with a decline in critical thinking

performance over time, suggesting that while automated summaries may maximize immediate retrieval efficiency, they may ultimately erode the ability of deep analytical capacity.

Future research should extend this design in three key directions to address the boundaries of these findings and the limitations inherent in the current design. First, the treatment must be replicated with longer, more conceptually demanding texts to test whether the observed AI summary advantage shrinks or reverses as the source material grows in length and inferential demand. Schools, in general, require analysis and comprehension of longer texts, and usually more complex literature rather than science fiction. By expanding the study in this direction, it would be possible to conclude findings more fit for use in real-world school settings. Then, future studies should measure recall over a longer longitudinal delay to distinguish a short-term retrieval advantage from durable learning, which would directly address the concern that AI summaries support immediate recall without supporting true knowledge transfer. Finally, to increase analytical precision, future designs should replace self-reported sleep metrics with objective measurements, such as uniform wearable tracking devices, to ensure the sleep variable carries greater data reliability. With these adjustments, the relatively distinct directional finding in this study—that an AI summary facilitates the immediate retrieval of explicitly stated themes on a short story without a measurable interaction with sleep—can be rigorously tested against the broader, more cautious picture painted by the literature on AI-assisted learning.

Works Cited

Alptekin, C., & Erçetin, G. (2011). Effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. TESOL Quarterly, 45(2), 235–266. https://doi.org/10.5054/tq.2011.247705

Baddeley, A. (1992). Working memory. Science, 255(5044), 556–559. https://doi.org/10.1126/science.1736359

College Board. (2025, October 6). New research: Majority of high school students use generative AI for schoolwork. https://newsroom.collegeboard.org/new-research-majority-high-school-students-use-generative-ai-schoolwork

Conte, F., Malloggi, S., De Rosa, O., Ficca, G., Righi, S., Viggiano, M. P., & Giganti, F. (2025). Sleep benefits prose memory consolidation in university students. Brain Sciences, 15(3), Article 265. https://doi.org/10.3390/brainsci15030265

Craik, F. I. M. (2002). Levels of processing: Past, present… and future? Memory, 10(5–6), 305–318. https://doi.org/10.1080/09658210244000135 A

Crowley, R., Alderman, E., Javadi, A.-H., & Tamminen, J. (2024). A systematic and meta-analytic review of the impact of sleep restriction on memory formation. Neuroscience & Biobehavioral Reviews, 167, Article 105929. https://doi.org/10.1016/j.neubiorev.2024.105929

Faber, M. (2023, September 27). Grant update: Impact of AI tools on reader comprehension. Duke Innovation Co-Lab. https://colab.duke.edu/blog-post/grant-update-impact-ai-tools-reader-comprehension/

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. Learning & Memory, 13(3), 259–262. https://doi.org/10.1101/lm.132106

Gavett, B. E., & Horwitz, J. E. (2011). Immediate list recall as a measure of short-term episodic memory: Insights from the serial position effect and item response theory. Archives of Clinical Neuropsychology, 27(2), 125–135. https://doi.org/10.1093/arclin/acr104

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

Newbury, C. R., Crowley, R., Rastle, K., & Tamminen, J. (2021). Sleep deprivation and memory: Meta-analytic reviews of studies on sleep deprivation before and after learning. Psychological Bulletin, 147(11), 1215–1240. https://doi.org/10.1037/bul0000348

Putri, F. R. (2025). A qualitative exploration of students' use of Humata.AI to understand thesis literature: A cognitive load perspective. Wiralodra English Journal, 9(2), 166–179. https://wej.unwir.ac.id/index.php/wej/article/view/471

Sridhar, S., Khamaj, A., & Asthana, M. K. (2023). Cognitive neuroscience perspective on memory: Overview and summary. Frontiers in Human Neuroscience, 17, Article 1217093. https://doi.org/10.3389/fnhum.2023.1217093

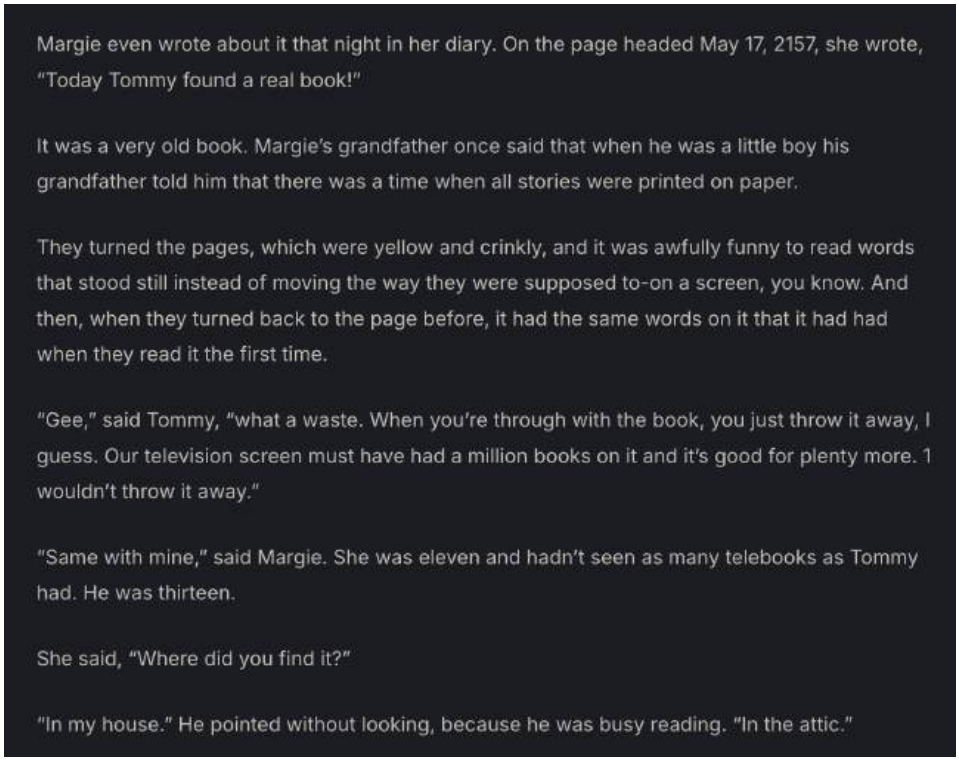
The Fun They Had Isaac Asimov (1951). (n.d.). https://www.johnspence.org.uk/wp-content/uploads/2022/11/The-Fun-They-Had.pdf

Vilcarino, J., & Langreo, L. (2025, October 8). Rising use of AI in schools comes with big downsides for students. Education Week. https://www.edweek.org/technology/rising-use-of-ai-in-schools-comes-with-big-downsides-for-students/2025/10

Yale School of Medicine. (2022, May 10). Sleep's crucial role in preserving memory. https://medicine.yale.edu/news-article/sleeps-crucial-role-in-preserving-memory/

Appendices

Full Text:



Margie even wrote about it that night in her diary. On the page headed May 17, 2157, she wrote, "Today Tommy found a real book!"

It was a very old book. Margie's grandfather once said that when he was a little boy his grandfather told him that there was a time when all stories were printed on paper.

They turned the pages, which were yellow and crinkly, and it was awfully funny to read words that stood still instead of moving the way they were supposed to-on a screen, you know. And then, when they turned back to the page before, it had the same words on it that it had had when they read it the first time.

"Gee," said Tommy, "what a waste. When you're through with the book, you just throw it away, I guess. Our television screen must have had a million books on it and it's good for plenty more. 1 wouldn't throw it away."

"Same with mine," said Margie. She was eleven and hadn't seen as many telebooks as Tommy had. He was thirteen.

She said, "Where did you find it?"

"In my house." He pointed without looking, because he was busy reading. "In the attic."

"What's it about?"

"School."

Margie was scornful. "School? What's there to write about school? I hate school."

Margie always hated school, but now she hated it more than ever. The mechanical teacher had been giving her test after test in geography and she had been doing worse and worse until her mother had shaken her head sorrowfully and sent for the County Inspector.

He was a round little man with a red face and a whole box of tools with dials and wires. He smiled at Margie and gave her an apple, then took the teacher apart. Margie had hoped he wouldn't know how to put it together again, but he knew how all right, and, after an hour or so, there it was again, large and black and ugly, with a big screen on which all the lessons were shown and the questions were asked. That wasn't so bad. The part Margie hated most was the slot where she had to put homework and test papers. She always had to write them out in a punch code they made her learn when she was six years old, and the mechanical teacher calculated the mark in no time.

"Well, he just told the boys and girls things and gave them homework and asked them questions."

"A man isn't smart enough."

"Sure he is. My father knows as much as my teacher."

"He can't. A man can't know as much as a teacher."

"He knows almost as much, I betcha."

Margie wasn't prepared to dispute that. She said, "I wouldn't want a strange man in my house to teach me."

Tommy screamed with laughter. "You don't know much, Margie. The teachers didn't live in the house. They had a special building and all the kids went there."

"And all the kids learned the same thing?"

"Sure, if they were the same age."

The Inspector had smiled after he was finished and patted Margie's head. He said to her mother, "It's not the little girl's fault, Mrs. Jones. I think the geography sector was geared a little too quick. Those things happen sometimes. I've slowed it up to an average ten-year level. Actually, the over-all pattern of her progress is quite satisfactory." And he patted Margie's head again.

Margie was disappointed. She had been hoping they would take the teacher away altogether. They had once taken Tommy's teacher away for nearly a month because the history sector had blanked out completely.

So she said to Tommy, "Why would anyone write about school?"

Tommy looked at her with very superior eyes. "Because it's not our kind of school, stupid. This is the old kind of school that they had hundreds and hundreds of years ago." He added loftily, pronouncing the word carefully, "Centuries ago."

Margie was hurt. "Well, I don't know what kind of school they had all that time ago." She read the book over his shoulder for a while, then said, "Anyway, they had a teacher."

"Sure they had a teacher, but it wasn't a regular teacher. It was a man."

"A man? How could a man be a teacher?"

"Well, he just told the boys and girls things and gave them homework and asked them questions."

"A man isn't smart enough."

"Sure he is. My father knows as much as my teacher."

"He can't. A man can't know as much as a teacher."

"He knows almost as much, I betcha."

Margie wasn't prepared to dispute that. She said, "I wouldn't want a strange man in my house to teach me."

Tommy screamed with laughter. "You don't know much, Margie. The teachers didn't live in the house. They had a special building and all the kids went there."

"And all the kids learned the same thing?"

"Sure, if they were the same age."

"But my mother says a teacher has to be adjusted to fit the mind of each boy and girl it teaches and that each kid has to be taught differently."

"Just the same they didn't do it that way then. If you don't like it, you don't have to read the book."

"I didn't say I didn't like it," Margie said quickly. She wanted to read about those funny schools.

They weren't even half-finished when Margie's mother called, "Margie! School!"

Margie looked up. "Not yet, Mamma."

"Now!" said Mrs. Jones. "And it's probably time for Tommy, too."

Margie said to Tommy, "Can I read the book some more with you after school?"

"Maybe," he said nonchalantly. He walked away whistling, the dusty old book tucked beneath his arm.

Margie went into the schoolroom. It was right next to her bedroom, and the mechanical teacher was on and waiting for her. It was always on at the same time every day except Saturday and Sunday, because her mother said little girls learned better if they learned at regular hours.

The screen was lit up, and it said: "Today's arithmetic lesson is on the addition of proper fractions. Please insert yesterday's homework in the proper slot."

Margie did so with a sigh. She was thinking about the old schools they had when her grandfather's grandfather was a little boy. All the kids from the whole neighborhood came, laughing and shouting in the schoolyard, sitting together in the schoolroom, going home together at the end of the day. They learned the same things, so they could help one another on the homework and talk about it.

And the teachers were people...

The mechanical teacher was flashing on the screen: "When we add the fractions 1/2 and l/4-"

Margie was thinking about how the kids must have loved it in the old days. She was thinking about the fun they had.

AI Summary

"The Fun They Had" takes place in the year 2157 and follows two children, Margie and Tommy, who discover a very old paper book about schools from centuries earlier. In their future world, children learn at home from mechanical teachers on screens that give lessons and grade homework automatically. Margie dislikes her robotic teacher, especially because it constantly gives her tests. When Tommy explains that children long ago went to school buildings with human teachers, Margie is surprised and curious. In those schools, kids learned together in classrooms, played in schoolyards, and helped each other with their work. After Tommy leaves, Margie returns to her mechanical teacher but keeps thinking about the old schools. She imagines how much fun it must have been to learn and spend time with other children, realizing that the old way of schooling might have been more enjoyable than her isolated, automated education.

System Design: Modern Large Language Models for Speech To Braille and Braille To Speech By Om Sandip Patel

Abstract

Studies show that Braille literacy is linked to an increase in academic and employment opportunities for visually impaired individuals. However, researchers have identified a beginning of the "Braille literacy crisis" which may reduce opportunities for students and adults. In this paper, we propose a system that serves as a complementary pathway to enhance education and inclusivity for visually impaired individuals. This system produces print-ready 3D models which will encode Braille cells with the correct geometry and spacing (dot diameters, heights, column/row spacing). The 3D STL file output enables the authoring of lectures or notes and requires only a 3D printer or any low-cost desktop printer capable of producing tactile pages, as compared to existing actuator-based Braille output systems which require expensive specialized equipment. This design also generates durable learning materials and provides better support for complex layouts (math, diagrams) without the need for expensive refreshable hardware. We designed two separate pipelines using open-source, cutting-edge innovations in Artificial Intelligence to convert Speech to 3D Braille and 3D Braille or BRL/BRF electronics files with braille content to Speech. We aim to better equip visually challenged individuals by bridging the gap between traditional learning methods and the ways visually-impaired individuals study.

Introduction

Braille is an important medium for visually impaired and blind students, improving their capabilities for communication and literacy. Studies show that Braille literacy is directly linked to an increase in academic and employment opportunities[49]. Braille makes it easier for students to study subjects which may be difficult to learn with the help of audio-only technologies, such as grammar, spelling, and reading comprehension[1]. Even though there are many benefits of Braille, studies show that the usage of Braille is declining, an event some researchers are calling a "Braille literacy crisis[2]. The current decline in Braille literacy makes it more important than ever to find better and newer ways to teach, learn, and collaborate in Braille for visually impaired students.

Available technologies underscore the shortcomings of existing approaches to Braille education. For example, the Vocal Braille system combines speech recognition with text-to-Braille conversion, sending Braille character by character to an optical character recognition (OCR) system, which is complemented with physical and audio feedback[50]. However, traditional refreshable Braille displays are often expensive and unavailable to the average person, and reading Braille character by character makes the system and user lose the context of the larger text. This makes scenarios such as sharing notes, authoring a book, or revising text difficult. These refreshable braille displays are useful for immediate information

[49] S. Hoskin, R. Greer, and D. McCarthy, Effectiveness of technology for Braille literacy education for children: a systematic review. 2022.

[50] C. A. Malik, K. T. Patil, K. S. Bhoir, and M. B. More, "Vocal Braille: Innovative speech-driven Braille display," Technische Sicherheit, vol. 25, no. 5, pp. 98–104, 2025.

exchange like conversation, but have limited utility for information storage and long form communication.

These limitations in existing technologies drove the motivation for this project and led to our research question: How might visually impaired students benefit from advancements in modern AI? We aimed to design a system that can improve Braille literacy and education, decreasing the barrier to entry in information sharing between visually impaired students, groups, and the broader community. By developing a strategy to turn speech into printed Braille, we will make it easier for individuals to share physical notes, author longer passages, or revise text. A Braille-to-speech pipeline could be helpful for instructors teaching students that cannot yet read Braille. By enabling easy translation of printed Braille to speech, this pipeline would enable students to combine more familiar auditory tools with written passages to help increase Braille literacy.

Here, we propose two minimum tolerance, low-latency pipelines that convert speech to 3D printable Braille and Braille documents to speech. Our system design also proposes a novel solution to real-world linguistic problems such as punctuation, contractions, numbers, and specialized notations. We have also addressed the constraints of Braille encoding systems, including Grade 1, Grade 2/Unified English Braille (UEB), and Nemeth code. This paper presents a practical but research-driven architecture for both Speech-to-Braille (S2B) and Braille-to-Speech (B2S) conversion, designed to improve literacy, accessibility, and independence for visually impaired students.

Methods

Our system design consists of two main pipelines, Speech-to-Braille (S2B) and Braille-to-Speech (B2S), outlined in figures 1 and 2 respectively. To develop our bidirectional S2B and B2S system, we utilized a range of open-source libraries, toolkits, and modern AI models that are available in Table 1.

Table 1

Library / ToolKit	Description and Usage	Download Links
OpenAI Whisper / Faster Whisper	A large, weakly supervised sequence-to-sequence model for speech recognition, offering multilingual support and robust accuracy. We referenced quantized GGML builds for efficient CPU deployment.	https://pypi.org/project/faster-whisper/
Kaldi / Vosk	Kaldi is a robust toolkit for streaming Automatic Speech Recognition (ASR), and Vosk provides lightweight offline models suitable for embedded boards like Raspberry Pi.	https://pypi.org/project/vosk/

Google Cloud Speech-to-Text	A cloud-based API with strong out-of-the-box accuracy and streaming features, including diarization and adaptation.	https://pypi.org/project/google-cloud-speech/
Mozilla Common Voice Dataset	A large, open dataset of speech samples with diverse accents and gender tags used for benchmarking ASR systems.	https://www.kaggle.com/datasets/mozillaorg/common-voice
Liblouis	An open-source Braille translator supporting UEB (Unified English Braille), Nemeth, and literary variants, with Python bindings (Pyliblouis, PyBraille) for integration.	https://liblouis.io/downloads/ brew install liblouis
OpenCV / scikit-image	Libraries for image preprocessing, including deskewing, illumination correction, and denoising for scanned Braille pages.	https://pypi.org/project/scikit-image/ https://pypi.org/project/opencv-python/
PyTorch / TensorFlow	Deep learning frameworks used to implement CNN classifiers for Braille dot detection in scanned images.	https://pypi.org/project/pytorch/ https://pypi.org/project/tensorflow/
OpenSCAD / FreeCAD / Mesh Libraries	Tools for generating parametric 3D models (STL files) of Braille pages with configurable metrics.	https://pypi.org/project/openscad-py/
FastSpeech2 / Transformer / HiFi-GAN / VITS	High-quality neural Text-to-Speech systems supporting SSML input, streaming synthesis, and voice control.	https://pypi.org/project/hifi-gan/ https://pypi.org/project/pyttsx3/

Table 1. Open source libraries used in our pipelines. The above table lists access links to download and use the open source libraries, tool kits and AI-models used in this project.

Speech-to-Braille (S2B) Pipeline

As outlined in Figure 1, we designed this end-to-end S2B conversion process in four steps:

1. **Speech to Text**
   We benchmarked candidate Speech to Text systems on the Mozilla Common Voice dataset[51] and other audio sets including lectures, noisy classrooms, and multiple languages. We evaluated three classes of Automatic Speech Recognition (ASR) systems:

[51] Mozilla, “Common Voice,” https://commonvoice.mozilla.org

Whisper family (OpenAI Whisper and faster-whisper), Kaldi/Vosk family, and Google Cloud Speech-to-Text.

2. **Punctuation & Normalization**
   This step of our system restores punctuation, and it segments sentences. This step expands dates and numbers according to locale rules. It also applies language-specific tokenization. Then, a Braille-aware contraction disambiguator determines whether to emit Grade 1 Unified English Braille (UEB) output, which provides a letter-for-letter translation of the text, or the more commonly used Grade 2 output, which uses contractions to shorten common words and letter combinations. The program also produces a per-token confidence score and an explanation decision trace for human review. This system has settings for language selection, custom words, and fallback policies that help to decide on ambiguous cases. This allows safe manual verification and downstream tooling.[52]

3. **Braille Translation**
   In this step, our system feeds the normalized text from the previous step into a Braille translator like Liblouis. Our pipeline system is using appropriate tables, Unified English Braille (UEB) or Nemeth, and we wrap it with a verifier. This verifier enforces the layout constraints like cells, lines, and page width. We add a validation step for contraction choices. Here we put a layer of rules engine to flag edge cases like ambiguous contractions, long numbers, inline math, or code. We have configurable settings to flag a document for manual review.

4. **Layout to Parametric 3D (STL)**
   The last step in our system pipeline is to turn the verified Braille text and page metadata into parametric 3D geometry (OpenSCAD/FreeCAD or small mesh library) that places standard cylindrical bumps on a base plate using configurable metrics (dot height ~0.6–0.9 mm, diameters, inter‑dot/inter‑cell spacing) and printer tolerances (nozzle, layer height). We now export STLs, offer print preview/per‑page settings, and run a printability validator that simulates fingertip probe checks (min distances, bump solidity) before slicing.

Braille-to-Speech (B2S) Pipeline

Description

As shown in Figure 2, we designed this pipeline in 4 steps. First, we convert a digital Braille format file to text, then we apply normalization and prosody mapping that takes care of

[52] R. Hoffmann (Ch. 14), “Speech, Text and Braille Conversion Technology” (survey of Braille conversion, actuators, TTS/ASR implications).

modern English abbreviation, math symbols, and data formats. And then finally we convert the text into speech so we can hear the braille document read aloud to us.

1. **Braille-to-Text Conversion**
   For Braille-to-Text conversion, our input can be either scanned braille document images or digital Braille file format document (BRL/BRF).
   For scanned Braille pages, the pipeline becomes an Optical Character Recognition (OCR) problem, especially for perceptible (touchable) dots. This step includes a high-resolution scan and preprocessing like deskew, illumination correction, and denoising prior to dot detection. For preprocessing, we have chosen the OpenCV/scikit-image library, and for dot detection, we are using small convoluted neural network (CNN) classifiers in PyTorch or TensorFlow for robust detection on noisy/aged paper. Next, our pipeline maps detected dot patterns to cell codes and passes cell codes to the translator, Liblouis. At the end of this step, we produce per-token confidence for dot detection. Whenever confidence is low, we run a verifier stage. This verifying process will be a round-trip conversion where we map detected cells to characters and reconstruct cells to compare. For engineering flexibility, we keep the OCR detector and the translator configuration based on modules so we can replace the dot detector, the translator backend (Liblouis), or the table set without altering the rest of the system.
   For electronic Braille called BRF or BRL, we utilize the Liblouis library, which is a validated and well-maintained translator, as the mapper from cells to textual characters. Liblouis is an open source engine with wide table coverage, including UEB, Nemeth, literary variants, and bindings for integration.[53] In our system design, we have settings or configs to evaluate both the core translator and the language and table support in Python using Pyliblouis or community wrappers like PyBraille or similar projects. As a result, we can call the translator from our pipeline and swap tables at runtime, and produce BRF/BRL-to-text output. [54]

2. **Normalization & Prosody Mapping**
   In this stage, our system normalizes text, by expanding numbers, dates, abbreviations, and converting structural cues into lightweight prosody annotations or Speech Synthesis Markup Language (SSML), which means we detect and add pauses, phrase breaks, emphasis, simple pitch or duration hints. We also attach confidence flags for any low-confidence passages and special tags marking areas that need extra notation so the downstream text-to-speech (TTS) system can produce more human-like speech.

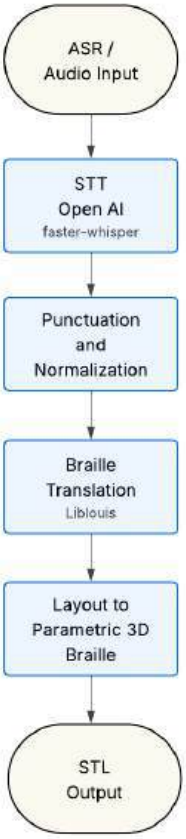
3. **Optional Cloud Processing**:
   We have added this step as optional, where we send only low-confidence or complex

[53] Liblouis: open-source Braille translator (https://liblouis.org).

[54] R. Hoffmann (Ch. 14), “Speech, Text and Braille Conversion Technology” (survey of Braille conversion, actuators, TTS/ASR implications).

segments like math, formulas, and long numerics to cloud models for disambiguation or advanced prosody. We have settings or configs for consent, encryption, and timeouts. We also provide a local fallback in case the network connection is not available. Cloud module returns normalized text/SSML and our design makes calls to be idempotent and parallelizable to avoid blocking real-time operation.

4. **Text-to-Speech**
   We use high-quality neural TTS for server-side rendering like FastSpeech2 / Transformer and front-end paired with HiFi-GAN or VITS to support streaming synthesis, SSML input, voice, and rate controls. FastSpeech2 is a high-quality, fast end-to-end TTS system.[55] This system cache renders audio and synthesizes segments for repeated texts and provides a robust fallback method for ultra-low confidence scores or failure cases.



[55] [2006.04558] FastSpeech 2: Fast and High-Quality End-to-End Text to Speech

Figure 2. Braille To Speech Pipeline, modules designed to make braille to speech in configuration driven architecture

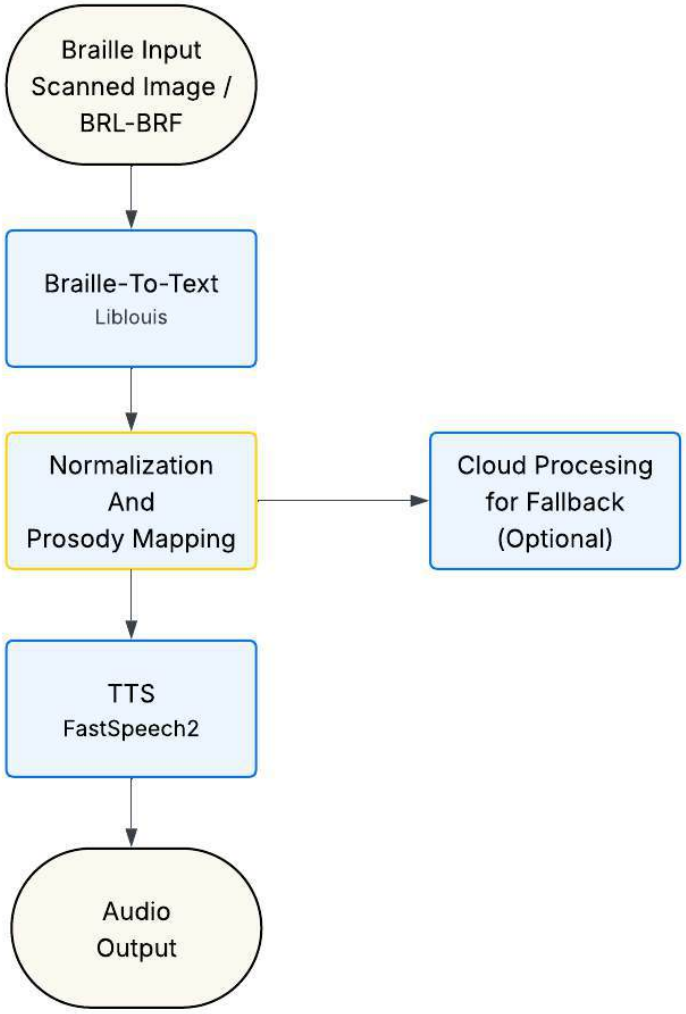


Figure 1. Speech to Braille Pipeline, modules designed to make braille to speech in configuration driven architecture

Results

Speech-to-Braille

In Speech-to-Braille conversion, the critical complex process is to convert first from speech to text. Research shows that Automatic Speech Recognition (ASR) has evolved through generations, from classical Gaussian Mixture Model (GMM), to Hidden Markov Model (HMM) and n-gram systems, to hybrid Deep Neural Network (DNN) acoustic models, to end-to-end sequence models. The recent deep acoustic models yielded large error-rate reductions in many benchmarks[56]. Connectionist approaches now enable alignment-free end-to-end mapping of audio to symbol sequences[57], and large end-to-end systems prove scalability on large written texts[58]. Multiple large language models for speech-to-text applications are publicly available. We ran an evaluation process and compared three different models for speed and accuracy. We evaluated Kaldi (and Kaldi-derived

[56] G. Hinton et al., "Deep neural networks for acoustic modeling in speech recognition," IEEE Signal Process. Mag., 2012.
[57] A. Graves et al., "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks," ICML, 2006.
[58] A. Hannun et al., "Deep Speech: scaling up end-to-end speech recognition," arXiv/2014 (Baidu DeepSpeech).

runtimes such as Vosk)[59], OpenAI Whisper (faster-whisper)[60], and Google's cloud based Speech-to-Text.[61]

Faster-Whisper is a large, weakly supervised, sequence-to-sequence model, with multiple language support and strong zero-to-few-shot robustness. We noted that the autoregressive decoder gives excellent accuracy and formatting with multilingual coverage. However, it is typically heavier than Connectionist Temporal Classification (CTC) models used in previous similar software. OpenAI Whisper can also be slower for a large-batch server without optimized runtimes.[62,63] We referenced quantized Georgi Gerganov's Machine Learning (GGML) builds that were tested for CPU deployment (int8/int4 tradeoffs).[64]

The Kaldi/Vosk family is a robust toolkit for streaming. It has low-footprint and Automatic Speech Recognition recipes. Vosk provides lightweight offline models with grammar and hint support. It can also run on embedded boards like Raspberry Pi for constrained domains and command-and-control use.[6566] But this family of systems generally lags large end-to-end DNNs on unconstrained conversational speech. Unlike more complex DNNs, Vosk simply transcribes word by word, not using the context of the conversation to add punctuation like question marks or recognizing a date and then writing that date according to your machine's settings.

We also evaluated Google Cloud Speech-to-Text as a representative of modern cloud-based technology. Cloud services provide strong out-of-the-box accuracy and streaming APIs that have diarization and adaptation features. However, they require network connectivity and raise privacy and cost considerations due to audio upload, retention policies, and per-minute pricing.[67]

After initial STT conversion, we added punctuation and normalization to add full stops, other punctuation, and math symbols. We then use a simple table to map each English character to its corresponding Braille character. As part of this system design for S2B conversion, we evaluated our pipelines using the Mozilla Common Voice dataset,[68] comprising 6956 test audio clips with diverse accent and gender tags.

Faster-Whisper generated speech to text with 87% accuracy at an average speed of 81 milliseconds (ms) per word, as shown in Figure 3. However, the Vosk model demonstrated 78% accuracy at 21ms per word. While the Vosk model is faster, the model does not accommodate

---

[59] D. Povey et al., "The Kaldi speech recognition toolkit," ASRU Workshop, 2011. (basis for Vosk runtimes)

[60] A. Radford et al. (OpenAI), "Robust Speech Recognition via Large-Scale Weak Supervision" (Whisper), 2022; faster-whisper implementations available for low-latency/offline use.

[61] Google Cloud, "Speech-to-Text Documentation," https://cloud.google.com/speech-to-text/docs

[62] A. Radford et al. (OpenAI), "Robust Speech Recognition via Large-Scale Weak Supervision" (Whisper), 2022; faster-whisper implementations available for low-latency/offline use.

[63] G. Klein, "faster-whisper," GitHub repo, https://github.com/guillaumekln/faster-whisper

[64] G. Gerganov, "whisper.cpp / ggml," GitHub repo (CPU quantized runtime), https://github.com/ggerganov/whisper.cpp ; https://github.com/ggerganov/ggml

[65] D. Povey et al., "The Kaldi Speech Recognition Toolkit," ASRU 2011.

[66] "Vosk API," alphacep GitHub, https://github.com/alphacep/vosk-api

[67] Google Cloud, "Speech-to-Text Documentation," https://cloud.google.com/speech-to-text/docs

[68] Mozilla, "Common Voice," https://commonvoice.mozilla.org

multiple language support or capitalization, question marks, or other prosody of English language. Google API produced 90% accuracy but it required 213ms per word as the model makes API calls to Google server. This model also needs internet connection which could violate privacy. Additionally, this service is not free, with per-minute charges for use. Detailed test results for each of our three models is available on GitHub, with the link included in Table 3.

We chose Faster-Whisper for its accuracy, open-source (free for public use) availability, and offline operation. Once we determined which model to use, we determined its robustness in different speech accents. Figure 4 shows the distribution of accents in the Mozilla dataset. Figure 5 displays the percent of words in each accent that were transcribed incorrectly by our finalized pipeline based on accent. The Malaysia and Philippines accents have the most incorrect words transcribed. The US and Canada accent came up with the best accuracy results. Google API comes with most accurate results, but requires much more time as the request goes to the Google server as an API call. That means using Google API in our pipeline will make internet connection a must have requirement. Also, Google API is not free, but is a paid service. We tested this model as a bench mark against our other free available options. Next best accuracy came from the Faster-Whisper model, which is free and does not require internet connection. Faster-Whisper took only 38% of the time of the Google Cloud model and has features like multi-language support, capitalization, and sentence formation that do not come with Vosk. These pros and cons led us to choose Faster-Whisper as the foundation for our S2B pipeline.

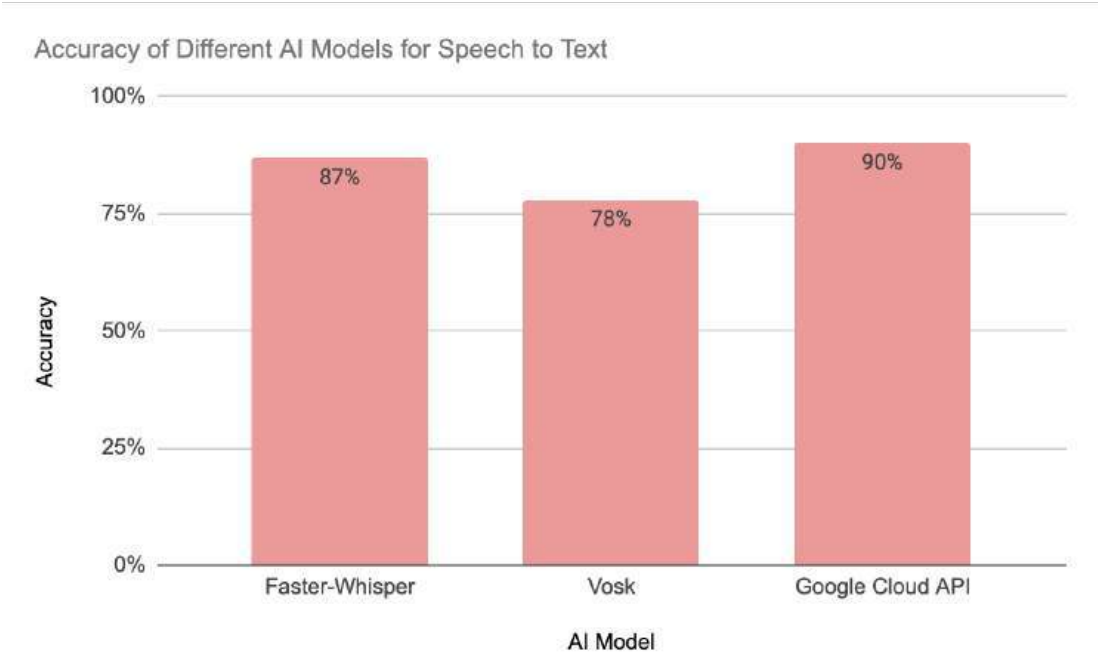


Figure 3: Accuracy of speech to text, Accuracy was tested by running a Mozilla data set of audio samples. Data is represented as percent of words that were accurately transcribed.

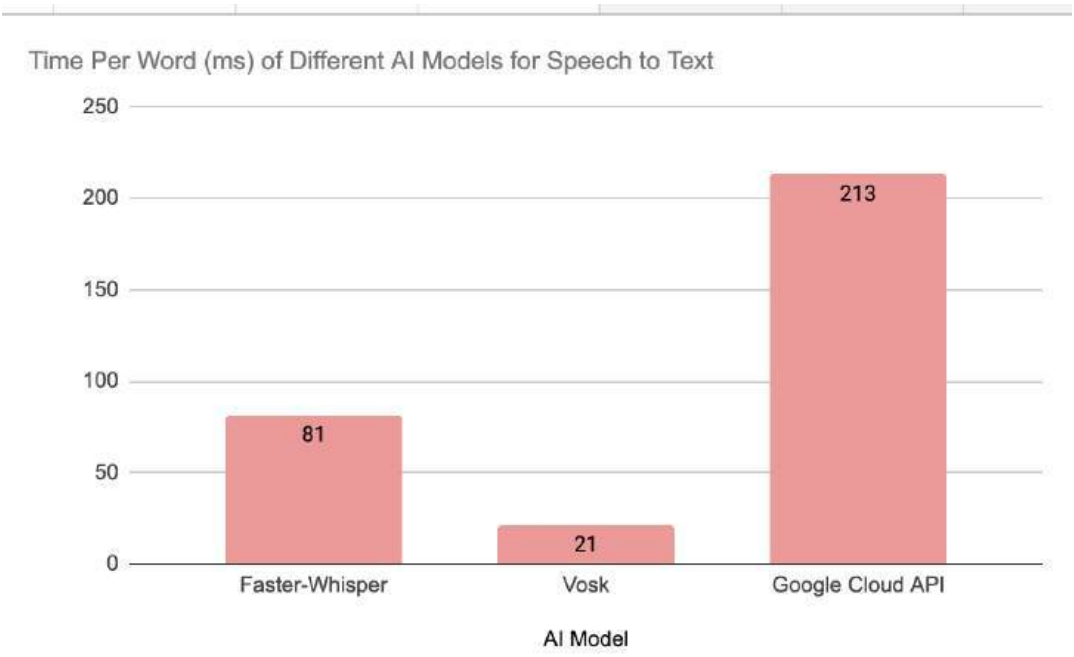


Figure 3.1: Time taken per word for speech to text Transcription speed was measured by running a Mozilla data set of audio samples. Data is represented as the time in milliseconds of how long the time took per word for different AI models.

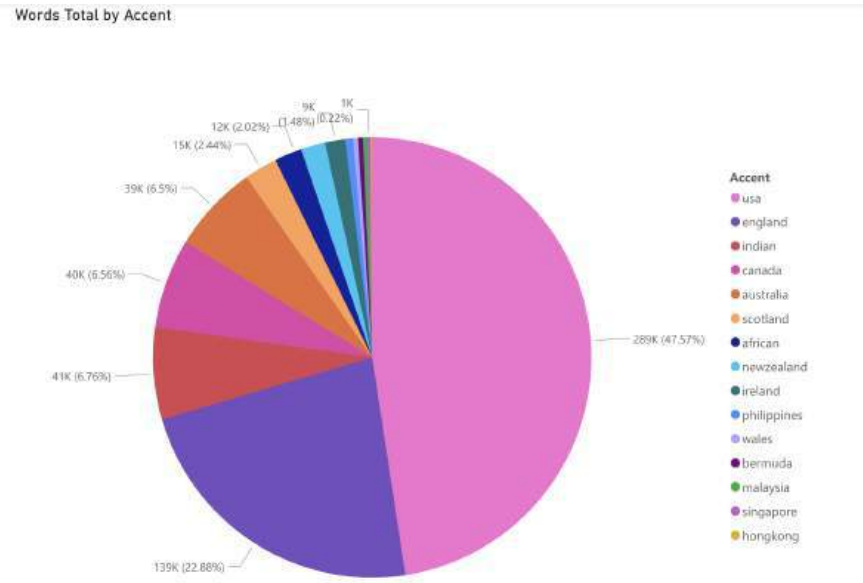


Figure 4: The Mozilla Common Voice dataset Our audio clip dataset included the accent of the speaker as a variable. This figure shows the data distribution of available audio clips by different accents.

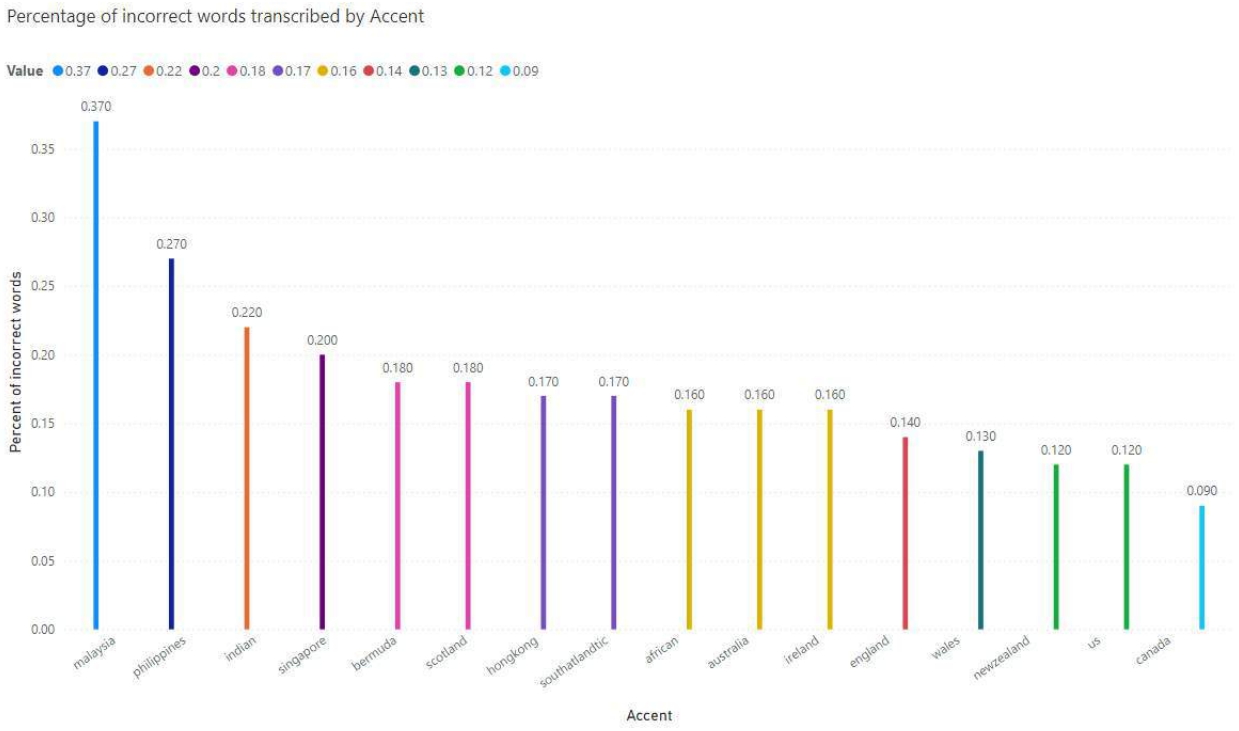


Figure 5: Percentage of incorrect words transcribed by accent. We tested the accuracy of our pipeline when evaluating audio samples with different accents. This graph shows the percent (as a decimal value) of incorrectly transcribed words by accent.

Braille-to-Speech

To perform an initial test of our pipeline's ability to translate Braille characters into text for an audio reader, we performed multiple steps shown in Figure 6. Figure 6 shows the full round trip pipeline with steps we executed to test this end to end accuracy. When reverting Braille transcriptions to speech, our pipeline showed complete accuracy, with no further errors introduced during this phase. The test output for B2S is included in the GitHub repository, with the link included in Table 3.

Our Text-to-Speech module produced high-quality, natural-sounding speech from normalized text, with streaming support and fallback mechanisms for reliability. The Braille-to-Speech pipeline required less optimization due to well-established tools for turning text into speech. The overall system generated durable, print-ready 3D Braille learning materials, supporting tactile readability and printer compatibility.

The most intensive portion of the B2S pipeline is the Braille-to-Text conversion. We used small CNN classifiers in PyTorch or TensorFlow for robust detection on noisy or aged paper. We tested our pipeline using scanned Braille for each character with different rotation, dim/bright lighting and front view vs side view orientation, a total 1561 pages from the Braille character dataset linked in Table 3 using our pipeline's dot detection process. This dataset includes scanned Braille characters in dim lighting, rotated images, and different whole side images. Our dot

detection algorithm is currently trained to recognize Braille images in dim lighting or regular lighting, but is not trained for rotated or side view images. Future work could enhance this capability. With all dim images, our system recognized Braille characters on scanned pages with 94% accuracy. Test results are drafted in Table 3. For Braille-to-Text conversion, our system reliably handled both electronic and scanned Braille documents, supporting multiple Braille tables and robust detection on noisy or aged paper. Our modular design allows flexible swapping of dot detectors, translation backends, and tables without affecting the rest of the system.

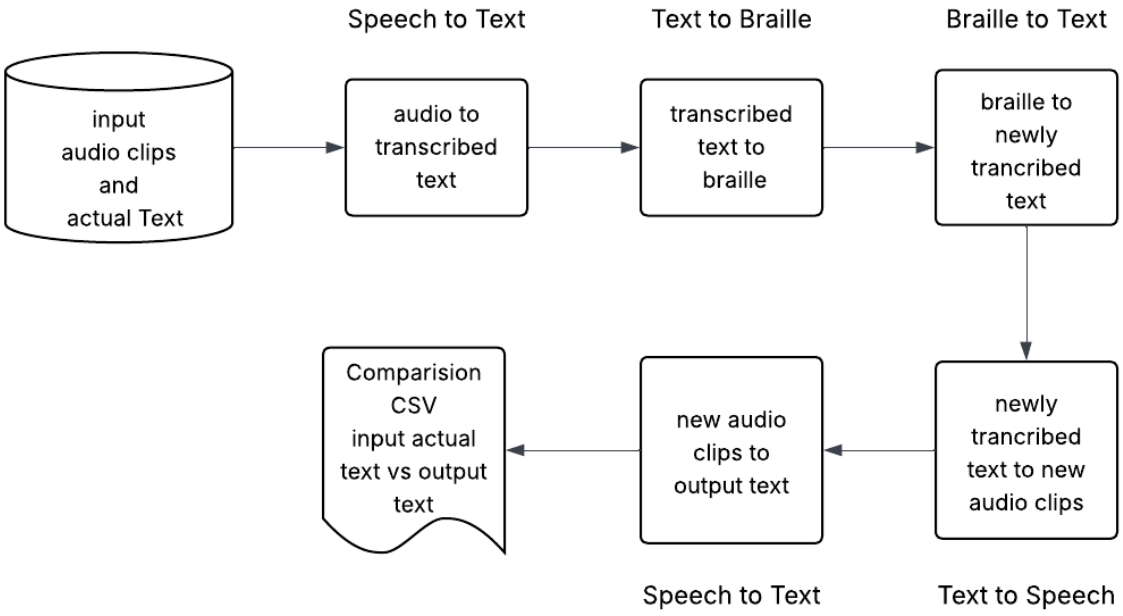


Figure 6. Roundtrip full pipeline, We used a full round trip pipeline to test Braille to speech accuracy. This diagram shows the steps taken in this pipeline.

Conclusion

We designed and implemented this system with a practical, modular architecture for bidirectional Speech↔Braille conversion. Prior work has shown typical hardware focus on electromechanical refreshable cells (solenoids, piezo pins, servos) that render characters cell-by-cell in real time. These devices prioritize immediacy and small footprint but are costly or are limited in reliability/rate (simple servos).[69,70,71] They do not generate persistent, page-level tactile artifacts suitable for low-cost fabrication. Thus, these devices do not function to author

[69] C. A. Malik, K. T. Patil, K. S. Bhoir, and M. B. More, "Vocal Braille: Innovative speech-driven Braille display," Technische Sicherheit, vol. 25, no. 5, pp. 98–104, 2025.
[70] G. Klein, "faster-whisper," GitHub repo, https://github.com/guillaumekln/faster-whisper
[71] G. Gerganov, "whisper.cpp / ggml," GitHub repo (CPU quantized runtime), https://github.com/ggerganov/whisper.cpp ; https://github.com/ggerganov/ggml

long communications or create sharable artifacts without losing its context as they render character by character in a STT scenario. Our pipeline aimed to provide a Speech-to-Braille strategy that would help produce long-format, long-lasting transcriptions.

Our S2B system stitches together modern Automatic Speech Recognition libraries, Braille-aware text normalization and contraction disambiguation, and a canonical Braille translator. We chose an optimized offline ASR system, Faster-Whisper, for its latency, accuracy, and privacy. S2B then feeds its output to a normalization stage that restores punctuation, expands numerics, and decides when to use Grade 1 or Grade 2 Braille transcriptions. As an output, this system generates durable learning materials that can be printed with any 3D printer. S2B uses a parametric 3D generator that yields printable, standards-compliant text pages.

Our B2S system can process electronic Braille input or camera-based scanned pages. B2S uses OpenCV/scikit-image, PyTorch, and Liblouis libraries to detect embossed dots, map them to Braille cells, and feed the normalized processed text to FastSpeech2 for speech synthesis. Our system cannot offer immediate conversational output, but rather is designed to create persistent output in physical format such as 3D printed Braille documents. We designed this system using engineering practices that emphasize verifiability and modularity. Our system uses a proven translator, Liblouis, wrapped with layout and round-trip checks. It provides a per-token confidence score and gives rationale for decisions to do manual review. We designed B2S to keep the OCR/detector, translator, and layout modules configurable at run time. This modular design makes it possible to switch tables or grades and detectors per locale and per user needs.

Our bidirectional S2B and B2S system balances constraints like latency, accuracy, cost, and privacy by supporting on-device offline applications, optional cloud disambiguation for low-confidence segments, and a unit-test harness of representative cases to catch contraction/layout regressions. We successfully can convert Speech to Braille and Braille to Speech. By improving this system, we can help visually impaired students author notes, books, and lectures, increasing opportunities for inclusivity in the visually impaired community.

Table 3

Title / Description	Reference links
Pipelines Code	https://github.com/omplays3889/speech-braille-speech
Speech to Braille Performance Test Results	https://github.com/omplays3889/speech-braille-speech/blob/main/Tests/SpeechToTextPipeline/performance-results-speech-to-text.csv
Braille to Speech Performance Test Results	https://github.com/omplays3889/speech-braille-speech/blob/main/Tests/BrailleToSpeechPipeline/performance-results-braille-to-speech.csv
Scanned Braille Performance Test Results	https://github.com/omplays3889/speech-braille-speech/blob/main/Tests/BrailleToSpeechPipeline/scanned_braille_results_updated.csv

Mozilla-Comon-Voice Data Set	https://www.kaggle.com/datasets/mozillaorg/common-voice
Scanned Braille Pages Data Set	https://www.kaggle.com/datasets/shanks0465/braille-character-dataset

Table 3. Code, performance test results and dataset for our pipelines, These are the links, where we implemented the system designed in this paper. You can refer to github links or csv files that are drafted as performance test results.

Glossary

ASR — Automatic Speech Recognition
STT — Speech-to-Text
S2B — Speech-to-Braille
B2S — Braille-to-Speech
NLP — Natural Language Processing
UEB — Unified English Braille
Nemeth — Nemeth Braille Code for Mathematics and Science Notation
HMM — Hidden Markov Model
GMM — Gaussian Mixture Model
DNN — Deep Neural Network
CTC — Connectionist Temporal Classification
OCR — Optical Character Recognition
BRF — Braille Ready Format (electronic Braille file format)
BRL — Braille file/format (generic use; specific meanings vary by toolchain)
WER — Word Error Rate
DNS — Deep Noise Suppression
GPU — Graphics Processing Unit
CPU — Central Processing Unit
STL — Stereolithography (3D model file format)
SSML — Speech Synthesis Markup Language
TTS — Text-to-Speech
HiFi-GAN — High-Fidelity Generative Adversarial Network (neural vocoder)
VITS — (end-to-end TTS architecture name; see VITS paper for exact naming)
LM — Language Model
API — Application Programming Interface
LID — Language Identification

Works Cited

S. Hoskin, R. Greer, and D. McCarthy, Effectiveness of technology for Braille literacy education for children: a systematic review. 2022.

[2] S. Hoskin, R. Greer, and D. McCarthy, Effectiveness of technology for Braille literacy education for children: a systematic review. 2022.

[3] C. A. Malik, K. T. Patil, K. S. Bhoir, and M. B. More, "Vocal Braille: Innovative speech-driven Braille display," Technische Sicherheit, vol. 25, no. 5, pp. 98–104, 2025.

[4] G. Hinton et al., "Deep neural networks for acoustic modeling in speech recognition," IEEE Signal Process. Mag., 2012.

[5] A. Graves et al., "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks," ICML, 2006.

[6] A. Hannun et al., "Deep Speech: scaling up end-to-end speech recognition," arXiv/2014 (Baidu DeepSpeech).

[7] D. Povey et al., "The Kaldi speech recognition toolkit," ASRU Workshop, 2011. (basis for Vosk runtimes)

[8] A. Radford et al. (OpenAI), "Robust Speech Recognition via Large-Scale Weak Supervision" (Whisper), 2022; faster-whisper implementations available for low-latency/offline use.

[9] C. A. Malik, K. T. Patil, K. S. Bhoir, and M. B. More, "Vocal Braille: Innovative speech-driven Braille display," Technische Sicherheit, vol. 25, no. 5, pp. 98–104, 2025.

[10] G. Klein, "faster-whisper," GitHub repo, https://github.com/guillaumekln/faster-whisper

[11] G. Gerganov, "whisper.cpp / ggml," GitHub repo (CPU quantized runtime), https://github.com/ggerganov/whisper.cpp ; https://github.com/ggerganov/ggml

[12] Mozilla, "Common Voice," https://commonvoice.mozilla.org

[13] R. Hoffmann (Ch. 14), "Speech, Text and Braille Conversion Technology" (survey of Braille conversion, actuators, TTS/ASR implications).

[14] Liblouis: open-source Braille translator (https://liblouis.org).

[15] R. Hoffmann (Ch. 14), "Speech, Text and Braille Conversion Technology" (survey of Braille conversion, actuators, TTS/ASR implications).

[16] [2006.04558] FastSpeech 2: Fast and High-Quality End-to-End Text to Speech

[17] Mozilla, "Common Voice," https://commonvoice.mozilla.org

[18] A. Radford et al. (OpenAI), "Robust Speech Recognition via Large-Scale Weak Supervision" (Whisper), 2022; faster-whisper implementations available for low-latency/offline use.

[19] G. Klein, "faster-whisper," GitHub repo, https://github.com/guillaumekln/faster-whisper

[20] G. Gerganov, "whisper.cpp / ggml," GitHub repo (CPU quantized runtime), https://github.com/ggerganov/whisper.cpp ; https://github.com/ggerganov/ggml

[21] D. Povey et al., "The Kaldi Speech Recognition Toolkit," ASRU 2011.

[22] "Vosk API," alphacep GitHub, https://github.com/alphacep/vosk-api

[23] Google Cloud, "Speech-to-Text Documentation," https://cloud.google.com/speech-to-text/docs

AI-Driven Marketing's Impact - Unpacking Measurable and Intangible Effects in B2B and B2C Contexts By Aarez Osman

Abstract

AI-driven Marketing impacts both business-to-business (B2B) and business-to-consumer (B2C) environments, focusing on measurable financial impacts and intangible effects such as customer loyalty, trust, and leadership. Through an extensive review of case studies and existing research, this study examines how AI enables personalization, predictive analytics, and cost efficiency in B2C markets, while affecting customer relationship management (CRM), sales forecasting, and complex decision-making in B2B markets. While both models pursue similar goals, such as revenue generation and efficiency, B2B and B2C companies differ in their strategies. B2B marketers rely on AI to obtain data-driven insights and improve long-term customer relationships. This contrasts with B2C companies, which place a higher focus on hyperpersonalization and customer engagement. In both cases, AI adoption has been shown to reduce costs, improve forecasting accuracy, and enhance customer satisfaction. However, some issues persist in AI adoption, such as algorithmic bias, regulatory compliance, and the "black box" problem. This research bridges gaps in B2B focused literature and hopes for future studies to develop more complex frameworks for measuring ROI, understanding ethical implications, and defining the evolving role of marketing professionals. Overall, this paper illustrates how AI-driven marketing is reshaping competitive strategies throughout the field, with distinct but complementary applications in B2B and B2C sectors.

Key Words

- Artificial Intelligence
- Digital Marketing
- B2B Marketing
- B2C Marketing
- Customer Relationship Management
- Marketing Automation

Introduction

In the Business-to-Consumer market, marketing driven by Artificial Intelligence is a powerful tool for personalization, consumer engagement, and revenue growth. B2C marketing has been the focus of most research regarding AI-driven marketing, making it a well-covered topic compared to B2B. B2C companies have used AI to hyperpersonalize their marketing, use predictive insights, and increase cost efficiency. They emphasize customer relationships, trying to build loyalty, retention, and stronger emotional connections by using AI to tailor their strategies to the customers' needs. Although concerns such as data privacy, algorithmic bias, and over-reliance on automation exist, AI's potential in reshaping how companies engage with buyers is vast. AI enhances day-to-day interactions and drives innovative strategies that deliver

an advantage in competitive markets. Overall, AI in B2C has reshaped customer experiences and delivered countless benefits for companies.

While existing literature has extensively discussed the application of AI-driven marketing in B2C companies, we know less about the possibilities of AI in B2B marketing and the comparison between the AI adoptions for the two. Similar to the B2C market, the integration of AI-driven marketing in B2B can offer enhanced opportunities to improve customer relationships and efficiency. Both sales and customer engagement can be improved through AI's data-driven insights, and although the human-centric nature of business relationships has posed an issue due to increased automation, countless benefits are poised to shape the future of B2B marketing. AI marketing in B2B has the opportunity to create new and innovative strategies that can deliver a competitive advantage to businesses.

While AI's impact in B2C markets focuses on individual customers, AI is used for more complex and larger sales cycles in a B2B market (Nystrom). AI has been shown to impact many aspects of B2B marketing, such as Customer Relationship Management (CRM), sales forecasting, value co-creation, operations management, and developing conceptual frameworks (Han et al.). The data-driven outlook that AI can help has greatly improved the customer relationships between businesses, and this data has been shown to help companies address any concerns clients may have before any escalation (Rajesh). Though AI has been proven to deliver benefits and advancements to marketing strategies, it has shown difficulty in adapting to more complex scenarios where human judgment is required (Keegan et al.). Occasionally, AI may oversell the benefits that its systems can deliver, leading to low satisfaction, which can hinder relationships between businesses (Keegan et al.).

This research provides in-depth insights through a literature review and compares the similarities and differences between B2B and B2C, as the existing research mainly focuses on the theoretical and broad structure of the topic, while not including a more detailed comparison of specific companies' strategies and outcomes. While B2C AI integration prioritizes short-term consumer engagement and rapid conversions, B2B frameworks leverage predictive analysis to navigate multi-stage pipelines and nurture long-term corporate relationships, realizing a distinct divergence in operational strategies despite shared financial goals.

The following section will start with a review of the literature on the impacts of AI on company performance, consumer satisfaction, and marketing leadership, followed by a discussion on future research areas and a conclusion.

The impacts of AI on company performance in B2C and B2B marketing

Artificial Intelligence has emerged as a revolutionary technology already impacting the future of business. AI-driven marketing increases revenue and customer acquisition by combining personalization, predictive insights, and cost efficiency. The integration of AI-driven marketing can redefine how businesses achieve financial success. Thus, it is crucial to explain how technological upgrades can impact a company's overall profit margin and its ability to generate revenue by enhancing sales and profitability, while also leading to cost reductions (Afua

et al.).

The core of any marketing strategy is sales, and the ability to impact a company's bottom line by enhancing sales and profit. AI has been shown to increase the customer lifetime value of a company, meaning that businesses have seen an increase in customer retention due to AI's ability to create a tailor-made experience for each customer (Akter Rainy; George et al.). Since AI can learn and recognize human patterns effectively, it can target customers with certain products that match their interests, which not only increases adoption but also helps build loyalty to the brand, building an enhanced value proposition (Sharma et al.; Zouaouia). This enhanced recognition of customer patterns is an improved response to their needs, with sources claiming that "76% of customers expect companies to understand their needs", and AI's use of predictive analysis helps quickly identify potential customers and their behaviors (Nalini et al.). This personalized experience has led to an increase in online sales, shown by a recent sharp increase in online vehicle sales (Boussairi).

Tesla, for example, uses AI to deliver a personalized experience, which has contributed to its greater market position in the automotive industry. Overall, AI has shown its ability to impact sales productivity by 25% in some cases by using predictive analysis in CRM systems (Stone et al.). AI also manages vast amounts of data, and its insights help teams make informed decisions, resulting in 90% accuracy in sales forecasts in some cases ("10 AI-Powered CRM Tools").

AI's ability to effectively market to a target audience online has allowed it to compete with more costly solutions that only larger companies with greater budgets can afford. One source shows how AI-driven social media marketing in Saudi Arabia has led to an overall increase in the number of customers and customer bases, while also increasing profitability. Social Media can also be a powerful tool due to the personalized algorithm that provides users with preferred content, and this algorithm can help deliver personalized ads to users as well (Basri). On top of finding potential customers, AI can provide assistance and additional information as a support chatbot, which can help convert some potential customers into buyers, as the added support can help customers find products and deal with any concerns a buyer may have. Since AI is such an advanced and rapidly growing technology, Kwabena Abrokwah-Larbi's research has shown how new developments such as augmented reality can help build a brand's image, leading to more customers and sales (Abrokwah-Larbi and Awuku-Larbi).

Other than its role in revenue creation, AI-driven marketing can also significantly reduce costs by automating many processes and optimizing marketing efficiency. AI can also reduce costs by ensuring that added revenue is spent efficiently due to AI's use of predictive analysis (Akter Rainy). The predictive analysis is a major factor in advanced targeting and will help ensure that there is not much wasted budget on unlikely customers (Sharma et al; George et al.). AI's minimal cost can aid smaller businesses in effectively increasing market share percentage, as it can compete with larger companies with greater resources and budget. Besma Boussairi's case study on Tesla shows the company was able to use AI to fully streamline its marketing process, using digital showrooms to generate engagement and visibility while significantly

decreasing its marketing budget (Boussairi). AI is also more efficient as it uses automation to repeat various tasks, some of which may take humans more time to complete. On top of the efficiency, AI and machine learning as a service have made AI implementation more cost-effective and lower long-term expenses (Östberg). AI's efficiency and elimination of human error have made AI a desirable alternative that can improve data security and reduce costs related to mistakes (Nalini et al; Abrokwah-Larbi and Awuku-Larbi).

In a B2B context, AI has been able to greatly aid an educational materials publisher, Sage Publications, which used AI to draft book descriptions. Human workers used to manually create descriptions based on the preface, title, and author of a book, and these descriptions are typically added to campaign channels such as ads, email, and social media, and are a main part of their marketing. Now, using AI has helped them by creating descriptions in seconds with only a preface, author, and title, while also using keywords to optimize search engine results(Zenon III). The added efficiency has led to fewer human resources required to complete the job, a much shorter timeframe to draft a description, as well as added marketing benefits with AI's knowledge of search engine optimization and marketing campaign personalization. Another benefit is AI's 24/7 availability, meaning that it can respond to customer inquiries without the cost of an around-the-clock team(Basri). Hallam Zouaouia's case study of Ouedkniss showed that they could use AI to maintain a high level of search engine optimization, leading to the site being at the top of search results at no cost (Zouaouia). These complex tasks, which AI can now complete quickly for little to no cost, allow human workers to focus on more strategic activities. On the B2B side, Salesforce's Einstein agent has also led to great financial success, as it's able to receive and analyze customer feedback 24/7, leading to a more personalized plan for business customers while also functioning without the cost of an around-the-clock team to analyze and receive feedback("10 AI-Powered CRM Tools").

Both B2B and B2C marketing have the same financial goals, but often use different strategies to achieve success due to different resources and customers. For example, B2C uses personalization to create a tailor-made experience designed solely for the user by matching customers' interests with products and personalized ads. On the other hand, B2B businesses use a more strategic and data-driven approach to identify high-intent accounts and contacts within their ICP and create personalized business plans based on these signals. To create revenue, B2C companies use AI to drive sales conversions for individual products, while also attempting to increase profitability and customer engagement using cheaper AI alternatives in marketing. B2B companies use AI in less direct ways and instead try to improve sales productivity and sales forecasts by using predictive analysis, while focusing on improving their CRM. Lastly, B2C uses AI to increase efficiency through tasks such as personalized content and more targeted marketing, while also reducing human resources by using AI chatbots for customer support. B2B uses highly specific and complex procedures, such as Sage publications using AI to draft book descriptions, reducing the need for humans and shortening the timeframe for their customers.

The Impact of AI on Consumer Satisfaction

AI-driven marketing's intangible effects on customer engagement and behavior demonstrate how AI can influence trust, brand loyalty, and perception beyond just financial benefits. It does so through personalized influences on emotions and decision-making (Chowdhury et al.). Examining how AI can impact customer acquisition, customer retention, targeting, and more efficient use of marketing budget will help provide a clear understanding of AI's increasing role in modern business and innovative customer engagement tools (Rani and Singh).

AI's ability to generate micro-personalized offers enhances customer satisfaction and builds further loyalty (Akter Rainy). This increased engagement and higher-quality marketing help create a stronger brand image and positive emotional bonds (Vithaldas Shah). Additionally, AI tools like chatbots provide immediate support, reducing wait times and boosting convenience (Gantumur). Augmented and virtual reality technologies also offer immersive ways to interact with products, adding a new dimension of engagement (Gantumur). The overall transformation of customer engagement strategies is a core benefit of AI adoption(Sivaraman).

Despite these benefits, there are risks related to privacy, as many believe their personal information is vulnerable as AI learns about their behaviors (Gantumur). This can significantly impact customer trust, highlighting the importance of developing responsible data handling practices (Nagy & Hajdu). Customers are more likely to trust companies that are transparent about their AI security policies, so openness about AI use is crucial for positive engagement (B. Z. et al.). Technologies such as AI fraud detection and risk assessment help enhance customer trust and security. However, reliance on AI for decision-making introduces the risk that technology may not meet expectations, and because the buying process involves many steps, poor performance at any stage could lead to customer dissatisfaction. Nonetheless, AI-driven marketing has proven successful in boosting engagement, with Salesforce reporting a 38% increase in lead conversions and a 37% rise in customer satisfaction (Tadimarri et al.).

AI has fundamentally transformed how companies build and maintain customer relationships by enabling a more personalized, data-driven approach that caters to individual needs. For example, Spotify and Netflix use AI to deliver customized recommendations based on previous user preferences, fostering emotional bonds and increasing engagement and retention (Romanchuk). Netflix analyzes viewing habits, time spent on the platform, search history, and ratings to reduce the time users spend searching for something to watch, which helps keep viewers engaged and has saved approximately $1 billion by lowering churn rates (Suvorov). Similarly, Amazon recommends products closer to a user's location for quicker delivery, improving the overall shopping experience (Suvorov). AI also enhances customer service; banks like Bank of America, Lemonade, L'Oréal, Walmart, and Chewy employ AI agents for 24/7 assistance (Wood). Sephora's AI virtual artist visualizes how makeup will look on a customer, increasing satisfaction and reducing returns by 30% (Wood).

In B2B, AI has been equally successful in strengthening customer relationships and managing complex sales cycles while nurturing high-value accounts over time. Salesforce's Einstein analyzes CRM data to personalize recommendations for procurement managers,

speeding up purchase decisions. This has led to a 67% increase in pipeline and a 25% rise in conversion rates for Siemens, a manufacturing firm ("10 AI-Powered CRM Tools"). HubSpot's AI agent, Breeze, customizes content at scale and predicts sales outcomes, while Xeople, a user of Breeze, achieved a 73% customer satisfaction rate, freeing up human resources through AI-powered sales agents ("10 AI-Powered CRM Tools"). Overall, B2B companies leverage AI to handle complex sales, enhance customer success, and strengthen high-value relationships.

Both B2B and B2C companies utilize similar AI technologies such as chatbots and predictive analytics to boost operational efficiency, customer service, and retention. Personalization enables tailored offers, product suggestions, and emails that foster stronger customer bonds. AI's data insights help companies improve the customer experience by incorporating customer feedback and preferences. Automation in tasks like chatbots, email marketing, and social media management increases efficiency and allows staff to focus on more critical activities while providing customers with better support.

Despite similarities, the methods and focus differ between B2B and B2C. B2B typically uses AI to solve complex problems such as lead generation and data analysis, whereas B2C emphasizes personalized recommendations and enhancing consumer experiences (Chang; Riemersma). Their strategies also diverge: B2B focuses on longer sales cycles and building relationships with fewer clients, while B2C aims for quick sales and wider engagement through high-volume, personalized marketing efforts. B2B companies tend to use more generative AI, like knowledge management and detailed data analysis, because their relationships are more complex and multi-stage (Riemersma). AI adoption in B2C is more centered on engaging consumers creatively through marketing activities, with an emphasis on high-volume engagement (Chang). Overall, AI is used in B2B marketing with a focus on managing existing pipelines and improving relationships with current customers, while B2C marketers' focal point is on enhancing customer experience and acquisition.

The Impact of AI on Marketing Leadership

Businesses face the intangible impact of AI-driven marketing, as it requires strategic, ethical decisions, reorganizing tasks, and providing different roles to workers (Stone et al.). Also, developing a different leadership model that has integrated AI into the workforce is necessary. Often, AI can offer a great alternative to regular leadership as its decision-making ability can present more data-backed alternatives and also reduce human cognitive bias (Ljepava). AI has been shown to offer a better alternative to traditional leadership in some cases, with B2C company Netflix using AI to analyze massive amounts of data, allowing leaders to receive insights and view trends that traditional human teams would not be able to provide (Suvorov).

Ethics is also a major part of marketing leadership, as leaders' decisions on AI usage and implementation can heavily influence brand reputation and public image (B. Z. et al.). The overall reliance on intuition and risk is greatly reduced. The lack of human bias can result in less managerial intuition and more of a focus on real data and results, improving the success of managerial tasks (Akter Rainy). On top of greater success, AI can automate many tasks, allowing

marketing leadership to take a greater role in the creative and strategic parts of their duties. Freeing up marketers' time can lead to more effort and focus on these higher-level and creative tasks, allowing for a higher quality marketing campaign overall (Vithaldas Shah, B. Z. et al.). An example in a B2B environment is HubSpot, whose AI tools have been shown to automate tasks such as personalized email campaigns, giving more time for humans to work on strategic planning ("10 AI-Powered CRM Tools"). AI also provides a greater level of feedback and can provide marketers with information that can help "optimize their campaigns, enhance customer engagement and identify other areas of engagement" (Östberg). There is also a greater level of adaptability with AI-driven marketing, as AI learns and can adjust campaigns and strategies to improve its overall efficiency (Akter Rainy).

On the contrary, some issues come to leadership when AI is implemented in a marketing campaign. Challenges include navigating algorithm bias and complying with regulations such as the General Data Protection Regulation (GDPR), all while maintaining customer trust (B. Z. et al.). On top of regulations, many marketing leaders face the "Black Box" problem, where AI makes decisions that are hard to explain, leading to less transparency and more time spent understanding the AI and developing clearer AI models (Pasquale). Furthermore, the 'black box' problem poses challenges to marketing leadership accountability when automated systems make inexplicable decisions. As shown by Papagiannidis, this lack of transparency can cause friction between strategy leaders and technical teams, leading to lower satisfaction and potentially harming corporate relationships rather than building customer engagement (Papagiannidis et al.). At one time, Amazon's AI-driven recruiting tool was found to have a gender bias, where AI found female applicants to be less preferable, and they were often excluded from candidacy (Parsani). This can present some challenges for businesses, as AI implementation is already an expensive and difficult process, as high-level technical experience, as well as ensuring your algorithms are not biased, are both required to fully integrate AI into your marketing team.

AI-driven marketing for B2B and B2C shows a similar foundation and similar opportunities for efficiency, while introducing a new set of ethical and strategic challenges. B2B and B2C companies use AI's ability to analyze data to make better marketing decisions, instead of depending on human intuition. Also, AI is used to automate routine tasks, placing less pressure on marketing teams to complete these responsibilities and allowing for time to be spent on improving their marketing strategies and high-level duties. Both types of companies face challenges as well, with the need to regulate algorithmic bias, follow all AI regulations, and work past the many challenges that implementing this new technology can cause.

Even with the many similarities, the application of AI-driven technology is different in a B2B and B2C context. Typically, B2C companies are focused on analyzing large-scale data when developing marketing strategies to appeal to their huge variety of customers while still trying to provide an ultra-personalized experience. B2B's AI is normally used for task automation, which aids a business through the lengthy sales process between businesses, while using customer feedback and trends to foster a positive relationship between the businesses. Overall, B2C businesses target diverse audiences, with personalized marketing to drive

immediate results and quick transactions, while B2B businesses place a higher focus on data-driven decisions.

Future Research Directions

This paper has explored the financial and intangible impact that AI-driven marketing can have in a B2B and B2C context, revealing the similarities and differences in strategies, applications, and outcomes. It was noticed that B2C uses AI for personalization and immediate sales, with a goal of also increasing customer retention. However, B2B companies tend to focus on more complex sales cycles and customer relationship management. Overall, both companies found issues such as leadership and ethics play a role in AI adoption, but both B2B and B2C companies have a shared goal to build wealth and increase company efficiency. Based on this literature review, several areas are identified in the following section as promising directions for future research in this rapidly evolving field, offering valuable opportunities for further scholarly inquiry.

First of all, a more detailed comparison between B2B and B2C could be beneficial for further research, as current examples may be high-level in explaining each type of business, but an in-depth comparison of the two would be greatly beneficial. Researchers could discuss the similarities and differences in marketing innovation, promotion, advertising, and strategic decision-making.

Another possible research topic is identifying the difficulties that can arise in finding numerical measures of the outcomes of AI applications, especially in B2B businesses. Many case studies are filled with qualitative words, with less concrete, quantitative data, which will help emphasize a point. This research would be able to create a framework to measure the ROI on AI-driven marketing and would be a valuable tool for businesses that are considering adopting AI. Also, most research has been focused on the benefits of AI adoption in marketing, and discussions surrounding the challenges and ethical concerns of AI could be a major aid to companies that are unsure of the downsides of AI. Researchers could go in depth on issues such as “the black box” issue, algorithmic bias, how to successfully follow GDPR and CCPA regulations, as well as the many ethical concerns that can arise with AI. Examining how B2B and B2C companies can ensure data privacy and build customer trust would be an extremely valuable resource.

An additional helpful area of research could be identifying the future of a marketing professional in a world where AI is rapidly evolving. Explaining how AI can be used as a partner, not a replacement, and how AI excels at automated tasks, and can free up marketers to complete more creative, high-level work. Touching upon the required new skills that marketing professionals would need, such as prompt engineering, data literacy, and the ability to oversee AI creations. Understanding the valuable qualities that a good marketer might have would be important as well, such as creative thinking and emotional intelligence, which would be critical to a good research paper, as it would be invaluable to the millions of current and future marketers.

Ultimately, AI-driven marketing is carving out a role in revolutionizing efficiency and building customer value, paving the way for a new era of strategic marketing filled with opportunities for research and innovation.

Conclusion

This paper demonstrates that while AI-driven marketing is key in transforming B2B and B2C contexts in measurable and intangible ways, key differences in applications, strategies, and outcomes lead to a need for deeper research and reflection on the future. AI is the future, and it can drive efficiency, hyperpersonalize, and deliver a competitive advantage to a variety of businesses. However, to obtain AI's full potential, gaps in the literature must be addressed, such as the majority of the focus being on B2C marketing, and there are limited resources on B2B marketing strategies, and the differences between B2B and B2C. Overall, this paper found that AI can boost revenue, reduce costs, improve forecasting, and enhance efficiency in both B2B and B2C contexts. However, the strategies differ, with B2C relying on personalization and B2B emphasizing CRM. Also, brand loyalty, trust, and engagement have been shown to increase with the personalization and predictive analysis that AI provides. Lastly, it was noticed that marketing supports data-driven strategies and can free leaders for creative tasks, challenging them with issues such as ethics and algorithmic bias. The future of AI should focus on creating datasets that deliver measurable ROI while addressing ethical challenges and redefining marketers' roles as the field evolves. Ultimately, AI-driven marketing is emerging as a revolutionary tool that can reshape how businesses create value, build relationships, and compete in B2B and B2C markets.

Works Cited

abhinav.japesh@superagi. “10 AI-Powered CRM Tools You Need to Know: A Beginner’s Guide to Seamless Integration - SuperAGI.” SuperAGI, 2 July 2025, https://superagi.com/10-ai-powered-crm-tools-you-need-to-know-a-beginners-guide-to-seamless-integration/.

Abrokwah-Larbi, Kwabena, and Yaw Awuku-Larbi. “The Impact of Artificial Intelligence in Marketing on the Performance of Business Organizations: Evidence from SMEs in an Emerging Economy.” Journal of Entrepreneurship in Emerging Economies, vol. 16, no. 4, Feb. 2023.

Afua, Wilhelmina, et al. “Transforming Financial Planning with AI-Driven Analysis: A Review and Application Insights.” World Journal of Advanced Engineering Technology and Sciences, vol. 11, no. 1, Feb. 2024, pp. 240–57, https://doi.org/10.30574/wjaets.2024.11.1.0053.

Akter Rainy, Tahmina. “AI-Driven Marketing Analytics for Retail Strategy: A Systematic Review of Data-Backed Campaign Optimization.” International Journal of Scientific Interdisciplinary Research, vol. 6, no. 1, Mar. 2025.

B., Zaghmout, et al. “Navigating Digital Leadership: How UK-Based Marketing Leaders Are Redefining Brand Trust in an AI-Driven Era.” British Journal of Management and Marketing Studies, vol. 7, no. 4, Dec. 2024, pp. 126–42, https://doi.org/10.52589/bjmms-qyhv6nau.

Basri, Wael. “Examining the Impact of Artificial Intelligence (AI)-Assisted Social Media Marketing on the Performance of Small and Medium Enterprises: Toward Effective Business Management in the Saudi Arabian Context.” International Journal of Computational Intelligence Systems, vol. 13, no. 1, Jan. 2020, pp. 142–52, https://doi.org/10.2991/ijcis.d.200127.002.

Boussairi, Besma. “AI Marketing and Its Role in Enhancing E-Commerce within Economic Enterprises: An Analytical Study of Tesla’s Experience (2012–2024).” International Journal of Economic Performance, vol. 8, no. 1, June 2025.

Chang, Sam. “AI Strategies for B2B and B2C: A Comparative Study in Marketing.” Rapidleadspro.Com, 2025, https://rapidleadspro.com/ai-strategies-for-b2b-and-b2c/. Accessed 31 Aug. 2025.

Chowdhury, Soummya, et al. “Influence of AI Driven Digital Marketing on Consumer Purchase Intention: An Empirical Study.” Journal of Informatics Education and Research, vol. 4, no. 2, Apr. 2024, https://doi.org/10.52783/jier.v4i2.811.

George, Sajan M, et al. “Migration Letters Role of Artificial Intelligence in Marketing Strategies and Performance.” Migration Letters, vol. 21, no. S4, Centivens Institute of Innovative Research, Feb. 2024, pp. 1589–99, https://www.researchgate.net/publication/378184143_Migration_Letters_Role_of_Artificial_Intelligence_in_Marketing_Strategies_and_Performance.

Han, Runyue, et al. “Artificial Intelligence in Business-to-Business Marketing: A Bibliometric

Analysis of Current Research Status, Development and Future Directions." Industrial Management & Data Systems, vol. 121, no. 12, Aug. 2021, pp. 2467–97, https://doi.org/10.1108/imds-05-2021-0300.

Keegan, Brendan James, et al. "Implementing Artificial Intelligence in Traditional B2B Marketing Practices: An Activity Theory Perspective." Information Systems Frontiers, vol. 26, no. 3, May 2022, https://doi.org/10.1007/s10796-022-10294-1.

Khuslen Gantumur. "The Impact of Artificial Intelligence Marketing Offers on Online Shopping Behavior." European Scientific Journal, ESJ, vol. 37, 2020, pp. 1–1, https://eujournal.org/index.php/esj/article/view/18973. Accessed 14 July 2026.

Ljepava, Nikolina. "AI-Enabled Marketing Solutions in Marketing Decision Making: AI Application in Different Stages of Marketing Process." TEM Journal, vol. 11, no. 3, Aug. 2022, pp. 1308–15, https://doi.org/10.18421/tem113-40.

Mrs.K. Nalini, et al. "Impact of Artificial Intelligence (AI) on Marketing." Int. J. Of Aquatic Science, vol. 12, no. 2, June 2021, pp. 3159–67.

Nyström, Ella. The Use of Artificial Intelligence in B2B Marketing Processes. 2024.

Östberg, Valtteri. AI-DRIVEN MARKETING ANALYTICS A Case Study on Enhancing Marketing Performance Reporting in SMEs. 2024.

Papagiannidis, Emmanouil, et al. "Uncovering the Dark Side of AI-Based Decision-Making: A Case Study in a B2B Context." Industrial Marketing Management, vol. 115, 2023, pp. 253–65, https://doi.org/10.1016/j.indmarman.2023.10.003.

Parsani, Puran. "Case Study: How Amazon's AI Recruiting Tool 'Learnt' Gender Bias." Cut-the-Saas.Com, 2024, https://www.cut-the-saas.com/ai/case-study-how-amazons-ai-recruiting-tool-learnt-gender-bias.

Rajesh, Sivakami. "An Empirical Study on Data-Driven Marketing for Enabling Better CRM in B2B." ResearchGate, unknown, 4 July 2024, https://www.researchgate.net/publication/381962701_An_Empirical_Study_on_Data-Driven_Marketing_for_enabling_better_CRM_in_B2B.

Rani, Shikha, and Vikramjit Singh. The Influence of AI Technologies on Consumer's Buying Decisions. May 2025.

Riemersma, Frans. "How B2B and B2C Brands Adopt genAI — Same Tech, Different Strategies | MarTech." MarTech, 13 Mar. 2025, https://martech.org/how-b2b-and-b2c-brands-adopt-genai-same-tech-different-strategies/. Accessed 31 Aug. 2025.

Romanchuk, Jenny. "How B2C Companies Leverage AI Marketing [Examples & Tips]." Hubspot.Com, HubSpot, 12 Mar. 2024, https://blog.hubspot.com/marketing/ai-b2c-marketing.

Sharma, Kapil Kumar, et al. "AI-Driven Marketing: Transforming Sales Processes for Success in the Digital Age." Journal of Knowledge Learning and Science Technology ISSN 2959-6386 (Online), vol. 2, no. 2, Sept. 2023, pp. 250–60,

https://doi.org/10.60087/jklst.vol2.n2.p260.
Sivaraman, Subhashree. “AI in Marketing: The Transformation of Customer Engagement Strategies.” ResearchGate, unknown, 24 Nov. 2023, https://www.researchgate.net/publication/375925389_AI_IN_MARKETING_THE_TRANSFORMATION_OF_CUSTOMER_ENGAGEMENT_STRATEGIES?enrichId=rgreq-c56a71ed5ee2072dd01bd6242c5f05d7-XXX&enrichSource=Y292ZXJQYWdlOzM3NTkyNTM4OTtBUzoxMTQzMTI4MTIwNjk4MDIxMEAxNzAxMDUyNjgyMTU1&el=1_x_2&_esc=publicationCoverPdf. Accessed 11 Aug. 2025.
Stone, Merlin, et al. “Artificial Intelligence (AI) in Strategic Marketing Decision-Making: A Research Agenda.” The Bottom Line, vol. 33, no. 2, Apr. 2020, pp. 183–200.
Suvorov, Victor. “AI-Driven Personalization: Cases of YouTube, Netflix & Amazon.” Elinext, 15 Apr. 2025, https://www.elinext.com/solutions/ai/trends/ai-driven-personalized-content-recommendation/.
Tadimarri, Anish, et al. “AI-Powered Marketing: Transforming Consumer Engagement and Brand Growth.” International Journal for Multidisciplinary Research, vol. 6, no. 2, 2024.
Team DigitalDefynd. “Exceptional B2C Customer Service [5 Case Studies][2025].” DigitalDefynd, 23 Aug. 2024, https://digitaldefynd.com/IQ/b2c-customer-service-case-studies/.
Vithaldas Shah, Tanmaykumar. “Leadership in Digital Transformation: Enhancing Customer Value through AI-Driven Innovation in Financial Services Marketing.” International Journal of Science and Research Archive, vol. 15, no. 3, GSC Online Press, June 2025, pp. 618–27, https://doi.org/10.30574/ijsra.2025.15.3.1767.
Wood, Christina X. “5 AI Case Studies in Customer Service and Support.” VKTR.Com, 24 July 2024, https://www.vktr.com/ai-disruption/5-ai-case-studies-in-customer-service-and-support/.
Zenon III, Alton. “How Sage, A Global Media Publisher, Uses AI.” Jasper.Ai, 2024, https://www.jasper.ai/blog/ai-global-media-publishing.
Zouaouia, Hallam. “The Effect of AI-Driven Marketing on Business Performance: Case Study of OUEDKNISS Website.” Al-Riyada for Business Economics Journal, vol. 11, no. 1, Jan. 2025.

Co-inoculation of AMF and *Trichoderma* on Tomato and Basil Production

By Alicia Naggah

Abstract

Fungi are widely beneficial components of agricultural systems, forming complex networks in soil that promote crop resilience, enhance soil health, and foster long-term sustainability. Mutualistic fungi have measurable effects on tree growth and crop yields resulting from increased nutrient uptake and healthier soil structure throughout the growth season. However, the effects of arbuscular mycorrhizal fungi (AMF) and *Trichoderma*, especially dual-inoculation, on plant growth rates and agricultural yield are not yet well understood. Currently, many studies examine the influence of AMF and *Trichoderma* on crop production separately rather than exploring the impact of co-inoculation and their potential to provide synergistic benefits. To address this knowledge gap, this study focused on understanding the impact of co-inoculating AMF and *Trichoderma* on vegetable crops compared to individual applications of AMF and *Trichoderma* to experimentally determine which application method resulted in greater observed growth. In this trial, 20 tomato and basil plants were treated with AMF, *Trichoderma*, and AMF and *Trichoderma* combined, in contrast with the control (no inoculation). Growth parameters, including plant height, plant width, and the number of leaves, were measured weekly for three weeks. At the end of the experiment, the dry aboveground biomass of the plants was recorded, and the data was analyzed in R. The experiment showed that co-inoculation of AMF and *Trichoderma* generally resulted in higher growth for basil and tomatoes, but their specific effects varied according to plant species. Under co-inoculation treatment, basil plants had greater plant width and leaf production, while tomato plants had more dry biomass and higher plant height. As climate change accelerates abiotic stresses like ion toxicity and high salinity, the potential of AMF and *Trichoderma* to alleviate these issues affecting crop yield remains overlooked. Future research related to mutualistic fungi can help protect global crop yield while providing sustainable alternatives to harmful pesticides.

Introduction

Fungi play significant roles in shaping plant growth, health, and productivity, as they can act as either pathogens that reduce crop yield or as symbiotic partners with plant roots, thereby increasing the uptake of soil nutrients. Beneficial fungi, such as arbuscular mycorrhizal fungi (AMF), have been found to increase tomato yield by 30% (Ullah et al., 2024). This reveals the potential of fungal symbioses to reduce synthetic fertilizer usage that has not been fully explored. Understanding the relationship between fungi and plant species is essential to developing bio-pesticides and other methods to secure food production without risking damage to the ecosystem.

Arbuscular mycorrhizal fungi (AMF) are root-associated symbiont fungi in the phylum Glomeromycota that form symbiotic relationships with plant roots to improve nutrient absorption from soil, most notably phosphorus, while receiving carbon from plants. This relationship is

widely observed since AMF forms associations with approximately 72% of vascular plant species across various natural and agricultural ecosystems around the globe (Brundrett & Tedersoo, 2018). AMF can be considered an essential component of plant nutrition, growth, and development, but scientists have only identified a limited number of their species due to environmental and geographical factors creating hidden biodiversity (Davison et al., 2015). Additionally, Trichoderma species exist as soil fungi that live in the soil but typically inhabit the rhizosphere because they help plants to grow and establish biochemical responses that prime the plant's immune system (Yu et al., 2022). The Trichoderma species includes more than 460 strains worldwide due to the fungi naturally occurring alongside crops in forest or agricultural soils, where they promote root growth, suppress pathogens, and boost plant health through their direct and indirect effects (Tyśkiewicz et al., 2022; Zin & Noor, 2020). AMF and Trichoderma coexist in the same soils because both organisms can infect host plants, which include tomatoes, corn, cucumbers, and soybeans (Poveda et al., 2019).

In natural and agricultural soils, plants interact with multiple fungi simultaneously, making it important to analyze their combined effects. Co-inoculation with multiple beneficial fungi, such as AMF and Trichoderma species, can produce favorable effects on plant growth, depending on environmental conditions and compatibility (Yang et al., 2022). Dual inoculation of tomato seedlings with Trichoderma and AMF significantly enhanced multiple growth factors compared with a control because each contributes different benefits to the plants, like enhanced uptake of nutrients and phosphorus solubility due to AMF and *Trichoderma*, respectively. (Mwangi et al., 2011). By investigating these fungi-plant interactions, we can create more ideal symbiotic environments for plant growth and improvements in crop yields. Improved crop yields can help reduce food insecurity and have a significant impact on a global scale, as they can help reduce food insecurity, especially under increasing food demand (Mumah et al., 2025).

In this experimental study, I investigated the effects of single and combined fungal inoculations on tomato and basil plants grown under controlled environments. Plants were assigned to the following treatments: no inoculation (control), AMF inoculation, *Trichoderma harzianum* inoculation, and co-inoculation with AMF + *T. harzianum*. Plants were monitored through repeated growth measurements (height, width, number of leaves, etc.). Based on previous experimentation and evidence that AMF can improve tomato growth and crop yield, and that Trichoderma can induce higher photosynthetic and respiratory rates for growth (Shoresh et al., 2008), I hypothesize that co-inoculated plants (AMF + *T. harzianum*) will result in the greatest growth compared to the control and single-inoculated groups due to complementary benefits in plant performance.

Methods

Equipment and Experimental Setup

To prepare for the experimental setup, I purchased materials before the start of the experiment. A 0.01g digital scale was purchased from Amazon for measuring dry biomass. The

pots the plants were grown in were 1-gallon nursery pots with twelve drainage holes, measuring 6.5 inches in diameter and 6.7 inches in height. I ordered non-GMO cherry tomato seeds from the brand Instant Latch, and I purchased Miracle-Gro potting soil from Home Depot for all plants. I propagated basil to reduce genetic variation by taking 20 cuttings, approximately 4 inches long, from a homegrown mother plant using the same soil in a larger pot. I placed the cuttings in cups filled halfway with filtered water and kept them near a window without direct sunlight. I replaced the water every five days, and the roots formed after 20 days.

During this propagation period, on day 13, I sowed tomato seeds by planting four seeds per 1-gallon pot in 20 pots to ensure successful germination. The other 20 pots (1-gallon size) were used for rooting basil cuttings, which were transplanted on day 21 of rooting in water. I labeled each pot with masking tape and a permanent marker to indicate species and plant number (e.g., B10 represents basil plant number 10). The experiment was designed with four groups containing five plants from each treatment group per species. Plants numbered 1, 5, 9, 13, and 17 served as a control; Plants 2, 6, 10, 14, and 18 were treated with *Trichoderma*; Plants 3, 7, 11, 15, and 19 were treated with AMF; Plants 4, 8, 12, 16, and 20 were treated with AMF and *Trichoderma* combined.

After transplant, the plants were allowed to adapt to their new surroundings for 12 days before I applied fungal treatments. The *Trichoderma* inoculum was sourced from Arbico Organics (Mikro-Root microbial fertilizer), a dry powder containing *T. harzianum* and *T. viride* at a concentration of 2.0 x 10^7 Cfu/g. The AMF was purchased as Organic MycoRoots Root Enhancer, containing 500 spores per gram with a total microbial concentration of 6.60 x 10^8 Cfu/g. I administered AMF to the plants by mixing 1 teaspoon per gallon of tap water used for soil-drenching the plants assigned to the treatment. I applied *Trichoderma* by mixing one teaspoon (2.5g) of the product per gallon of tap water to soil-drench the designated plants. After inoculation, I monitored the plants for three weeks with weekly measurements of plant height, plant width, and the number of leaves recorded for each plant. Following each measurement, I re-randomized pot positions to reduce environmental bias that could impact results. The plants were placed outdoors in a backyard in New York City during August 2025, spaced four inches apart and arranged into five rows of eight plants, with their species and predetermined treatments randomized. During the growth period, the plants were exposed to natural environmental conditions, such as temperatures reaching 21 to 35 degrees Celsius, receiving 7 to 8 hours of sunlight daily, plus exposure to local weather patterns and insects. The plants were watered daily at 6:00 PM without any fertilizers or pesticides.

Starting from day 25 of outdoor growth, I measured plant height and the number of leaves before fungal application. I used a ruler to measure plant height and plant width with 0.1 cm accuracy. The recorded measurements, including the number of leaves, during transplant for each plant, were organized on a spreadsheet and recorded weekly. At the end of the experiment, I recorded the final measurements before harvesting above-ground tissue. I clipped the plants at the soil surface, gently brushed them to eliminate dirt, and placed them on parchment paper. Samples were labeled and placed in an oven set at ~52 degrees Celsius for 52 hours, where they

were completely dried. The dry biomass was measured to 0.01g accuracy with the digital scale, and each plant's measurements were recorded.

Statistical Analysis

R software was used to evaluate the effects of the fungi inoculation on plant growth. Treatment types (Control, *Trichoderma*, AMF, and AMF + *Trichoderma*) were the independent variables, with plant height, plant width, the number of leaves, and dry biomass as the dependent variables. I used the difference between the values at the end of the experiment and before the fungal treatment, and the difference between the values before harvesting and before fungal treatment, for data analysis. I used the recorded values for the dry biomass of the basil and tomato plants. The dplyr package, used for data manipulation, performed the analysis. Two separate experiments were recorded for the basil and tomato plants to account for different growth patterns in each plant species. The analysis of plant height and the number of leaves was conducted at five different times: before fungi treatment, 10 days after treatment, 17 days after treatment, at the end of the experiment, and before harvesting. The analysis of plant width includes measurements taken 10 days after treatment, 17 days after treatment, at the end of the experiment, and before harvesting. The dry biomass for both species was also evaluated after harvesting. One-way analysis of variance (ANOVA) tests were used to assess any statistical differences between treatment groups at each time frame across the dependent variables. Additionally, Tukey HSD tests were used as a post hoc analyzer when ANOVA results showed significant treatment effects to identify pairwise differences between treatments. Boxplots were created with the ggplot2 package to portray how the growth factors spread across different treatments and time points through data visualization.

Results

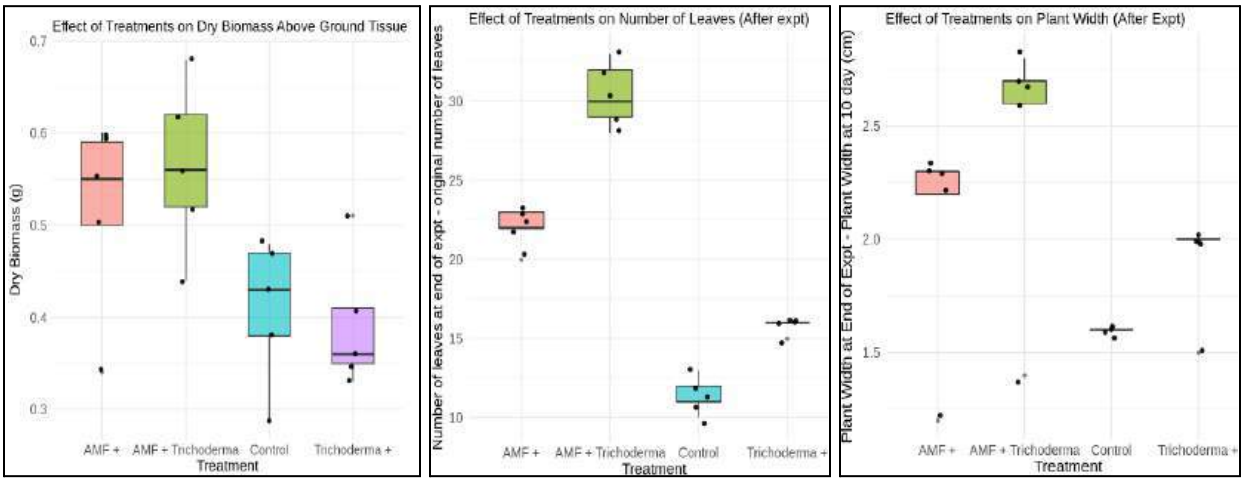
Experimental Results for Basil Plants

The application of fungal treatments influenced the growth characteristics of basil plants, which include their dry biomass, leaf production, and plant width. The dry biomass of above-ground tissue for basil plants showed significant differences across different treatments (one-way ANOVA, p-value = 0.0188). The Tukey HSD post hoc analysis showed that basil plants that received both AMF and *Trichoderma* treatment produced higher dry biomass than those with *Trichoderma* alone ($p < 0.05$), while the combined treatment showed no significant statistical difference when compared to both the control and AMF-only treatments (Fig. 1a).

The final number of basil leaves produced at the end of the experiment showed distinct differences between treatments (one-way ANOVA, $p < 0.001$). Tukey HSD analysis showed p-values less than 0.05, revealing that all treatment groups differed significantly from each other (Fig. 1b). Basil plants treated with AMF and *Trichoderma* produced the greatest number of leaves, with the control group producing the fewest. Single inoculation treatments resulted in

intermediate production, with AMF treatment producing higher leaf counts than the *Trichoderma* treatment.

At the halfway point of the experiment, plant width measurements showed no significant differences between the different treatments (one-way ANOVA, $p > 0.05$). When the study concluded, plant width measurements showed statistically significant differences between the various treatments ($p < 0.001$). The Tukey HSD tests revealed that basil plants treated with AMF and *Trichoderma* had greater plant width than all other treatment groups ($p < 0.05$). The single-inoculation treatments showed wider plant growth than the control group but did not differ significantly from each other (Fig. 1c).



(1a) (1b) (1c)

Figure 1: Effect of treatments on the number of leaves and plant width on basil plants at the end of the experiment (day 39) and dry biomass after harvesting (day 60). The thick line in the middle of each box represents the median value, the box represents the 75% quantile, the whiskers represent the 95% quantile, and the dots show the original values measured from each plant.

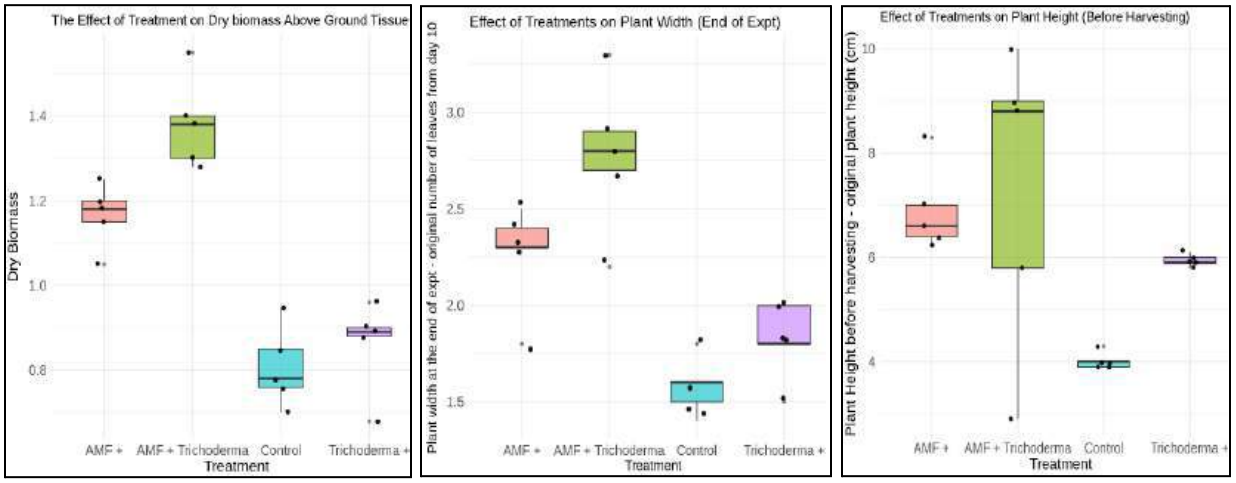
Experimental Results for Tomato Plants

Additionally, fungal treatments also had a strong effect on tomato biomass and plant structure. The above-ground dry biomass showed significant differences between treatments (One-way ANOVA, $p < 0.001$). The Tukey HSD analysis showed that tomato plants treated with AMF and *Trichoderma* had greater dry biomass than the other treatment groups ($p < 0.05$). Plants treated with only AMF produced more biomass than the control group, with the difference statistically significant ($p = 0.001$). The plants treated with *Trichoderma* alone showed no significant difference in dry biomass, as shown in the results (Fig. 2a).

The tomato plant width measurements at the end of the experiment showed significant treatment differences (One-way ANOVA, $p < 0.001$). The Tukey HSD analysis demonstrated that AMF and *Trichoderma*-treated plants had greater plant width ($p < 0.05$) than the control group and single-inoculation treatments. The plants treated with only AMF showed greater width

than the control plants, while plants treated with *Trichoderma* alone did not differ significantly from the control (Fig. 2b). In contrast to width and biomass, no significant differences were observed among treatments for the number of tomato leaves at the end of the experiment ($p < 0.05$).

The experimental results showed that tomato plants had more height fluctuations during the testing period. The height of tomato plants measured before harvesting showed statistically significant differences (One-way ANOVA, $p = 0.0163$). The Tukey HSD analysis reported that both plants treated with AMF and plants treated with both AMF and *Trichoderma* were significantly taller than control plants ($p < 0.05$). However, the two treatment groups showed no significant differences in their final plant height measurements ($p > 0.05$). The plants treated with *Trichoderma* showed no significant difference from the control group, as well as the AMF and combined treatment groups (Fig. 2c).



(2a) (2b) (2c)

Figure 2: The effect of treatments on the plant width of tomato plants at the end of the experiment (day 39) and plant height before harvesting and dry biomass (day 60). The thick line in the middle of each box represents the median value, the box represents the 75% quantile, the whiskers represent the 95% quantile, and the dots show the original values measured from each plant.

Discussion

Findings Interpretation

The study results confirmed the initial hypothesis that co-inoculating plants with AMF and *Trichoderma* yields greater growth than in single-inoculated and control plants. The combined treatment in basil showed significant improvements in leaf growth and overall plant development, specifically in plant width and leaf production by the end of the experiment. In tomato plants, the combined treatment resulted in the greatest above-ground dry biomass and plant width, indicating enhanced structural growth. The AMF and AMF + *Trichoderma* treatments produced better results than the control for tomato height and biomass measurements,

suggesting improved resource acquisition and allocation efficiency. However, all treatment groups in tomato plants showed no significant differences in leaf number. The results indicate that fungal inoculation, especially the combined fungi treatments with possible synergistic effects between AMF and *Trichoderma*, led to biomass growth and increased width rather than leaf production, which may reflect changes in root function and nutrient uptake rather than leaf formation. Together, these findings demonstrate that multi-fungal inoculation enhanced plant performance, despite the magnitude and type of growth responses varying between species.

Similar Findings

AMF extends root surface area through extraradical hyphae networks, allowing plants to absorb phosphorus and essential micronutrients beyond their root depletion zones (Smith & Read, 2008). Additionally, AMF colonization increases root absorptive capacity by causing the formation of lateral roots, resulting in greater root length and soil exploration (Smith & Smith, 2011; Brundrett & Tedersoo, 2018). *Trichoderma* simulates root branching and elongation through its auxin-like and ethylene-modulating effects to create better root system architecture for nutrient uptake (Contreras-Cornejo et al., 2009; Harman et al., 2004). Recent work also demonstrates that using AMF with *Trichoderma* alters rhizosphere microbial communities, resulting in better nutrient mineralization and improved plant uptake efficiency from organic matter (Zhang et al., 2026). Multifungal inoculations improve plant growth by increasing nutrient absorption, hormone signaling, rhizosphere processes, and soil nutrient cycling (Yang et al., 2022). The combined effects of AMF root surface expansion and Trichoderma root structure development likely contributed to the substantial growth increases observed in the plants that received both treatments, aligning with the findings.

Broader Implications

The broader implications of these findings can demonstrate their impact on both sustainable agricultural practices and climate-resilient cropping systems. Mycorrhizal networks are becoming more recognized as essential elements that govern soil carbon movement and ecosystem resilience, emphasizing their role as an under-protected component of climate restoration (Li et al., 2025). AMF produces glomalin and stabilizes soil aggregates to enhance carbon sequestration and soil structure, contributing to long-term carbon storage (Rillig & Mummey, 2006). New agricultural research shows that using AMF with beneficial fungi in crop fields reduces the need for fertilizers while improving nutrient cycling and minimizing environmental runoff, which helps establish environmentally friendly agricultural systems that are driven by natural processes (Chowdhary & Songachan, 2025). For example, soil communities of AMF established by winter green manure crops have been shown to boost rice yield quality and nutrient content through their increased AMF spore density and community structure (Cao et al., 2024). Moreover, combined AMF-*Trichoderma* systems can improve the nutrient release process from both organic residues and crop straw, which demonstrates their potential use in circular agricultural systems that recycle on-farm materials instead of using

synthetic fertilizers. The observed effects in this study support that fungal inoculants should be used as agricultural tools for increasing productivity while improving both soil health and carbon (Zhang et al., 2026). In the context of climate change, as stressors such as soil salinization and ion toxicity are amplified by elevated evaporation rates and seawater intrusion, agricultural systems that depend on coastal and arid regions may experience declining productivity, making the understanding of fungal symbiotic relationships increasingly important. Future studies are needed to investigate fungal relationships across different crop types, soil conditions, and abiotic stresses, while also considering interactions with native microbial communities.

Conclusion

As clearly demonstrated throughout this paper, co-inoculation with AMF and *Trichoderma* boosts structural development and biomass growth in tomato plants, while also improving growth parameters in basil. The extended root systems of AMF fungi and the increased rhizosphere activity of *Trichoderma* suggest that their complementary functions create beneficial synergies for agricultural systems. The usage of multiple fungi creates a sustainable agricultural system through their ability to enhance nutrient cycling and soil structure. However, it remains unclear whether *Trichoderma*, which is generally salt-tolerant, can buffer AMF against salt stress to preserve symbiotic benefits in degraded soil. If *Trichoderma* can successfully shield AMF from these harsh conditions, it would allow farmers to utilize the pairing even in areas experiencing severe drought or salinization. Interpreting beneficial fungal symbiosis interactions may create an effective method to protect worldwide crop production against accelerating environmental changes.

Works Cited

Brundrett, Mark C., and Leho Tedersoo. “Evolutionary History of Mycorrhizal Symbioses and Global Host Plant Diversity.” New Phytologist, vol. 220, no. 4, 22 Jan. 2018, pp. 1108–1115, https://doi.org/10.1111/nph.14976.

Contreras-Cornejo, Hexon Angel, et al. “Trichoderma Virens, a Plant Beneficial Fungus, Enhances Biomass Production and Promotes Lateral Root Growth through an Auxin-Dependent Mechanism in Arabidopsis.” Plant Physiology, vol. 149, no. 3, 28 Jan. 2009, pp. 1579–1592, https://doi.org/10.1104/pp.108.130369.

Davison, James, et al. “Global Assessment of Arbuscular Mycorrhizal Fungus Diversity Reveals Very Low Endemism.” Science, vol. 349, no. 6251, 2015. Science, https://doi.org/10.1126/science.aab1161.

Harman, Gary E., et al. “Trichoderma Species — Opportunistic, Avirulent Plant Symbionts.” Nature Reviews Microbiology, vol. 2, no. 1, Jan. 2004, pp. 43–56, https://doi.org/10.1038/nrmicro797.

Mumah, Edwin, et al. “Exploring the Reality of Global Food Insecurity and Policy Gaps.” Humanities and Social Sciences Communications, vol. 12, no. 1, Aug. 2025, doi:10.1057/s41599-025-05315.

Mwangi, Monda, et al. “(PDF) Inoculation of Tomato Seedlings with Trichoderma Harzianum and Arbuscular Mycorrhizal Fungi and Their Effect on Growth and Control of Wilt in Tomato Seedlings.” Research Gate, Apr. 2011, www.researchgate.net/publication/307683761_Inoculation_of_tomato_seedlings_with_Trichoderma_Harzianum_and_Arbuscular_Mycorrhizal_Fungi_and_their_effect_on_growth_and_control_of_wilt_in_tomato_seedlings.

Poveda, Jorge, et al. “Trichoderma Harzianum Favours the Access of Arbuscular Mycorrhizal Fungi to Non-Host Brassicaceae Roots and Increases Plant Productivity.” Scientific Reports, vol. 9, no. 1, 12 Aug. 2019, https://doi.org/10.1038/s41598-019-48269-z.

Shoresh M, Harman GE. The relationship between increased growth and resistance induced in plants by root colonizing microbes. Plant Signal Behav. 2008 Sep;3(9):737-9. doi: 10.4161/psb.3.9.6605. PMID: 19704844; PMCID: PMC2634575.

Smith, Sally E., and David Read. “Structure and Development of Ectomycorrhizal Roots.” Mycorrhizal Symbiosis, 2008, pp. 191–X, https://doi.org/10.1016/b978-012370526-6.50008-8.

Smith, Sally E., and F. Andrew Smith. “Roles of Arbuscular Mycorrhizas in Plant Nutrition and Growth: New Paradigms from Cellular to Ecosystem Scales.” Annual Review of Plant Biology, vol. 62, no. 1, 2 June 2011, pp. 227–250, https://doi.org/10.1146/annurev-arplant-042110-103846.

Tyśkiewicz, Renata, et al. “Trichoderma: The Current Status of Its Application in Agriculture for the Biocontrol of Fungal Phytopathogens and Stimulation of Plant Growth.” International Journal of Molecular Sciences, vol. 23, no. 4, 19 Feb. 2022, p. 2329,

www.ncbi.nlm.nih.gov/pmc/articles/PMC8875981/#:~:text=The%20biocontrol%20action%20of%20Trichoderma, https://doi.org/10.3390/ijms23042329.

Ullah, F., Ullah, H., Ishfaq, M. et al. Genotypic variation of tomato to AMF inoculation in improving growth, nutrient uptake, yield, and photosynthetic activity. Symbiosis 92, 111–124 (2024). https://doi.org/10.1007/s13199-023-00961-5

Yang, Rong, et al. "The interactions between arbuscular mycorrhizal fungi and Trichoderma longibrachiatum enhance maize growth and modulate root metabolome under increasing soil salinity." Microorganisms 10.5 (2022): 1042.

Yang, R., Qin, Z., Wang, J., Zhang, X., Xu, S., Zhao, W., & Huang, Z. (2022). The Interactions between Arbuscular Mycorrhizal Fungi and Trichoderma longibrachiatum Enhance Maize Growth and Modulate Root Metabolome under Increasing Soil Salinity. Microorganisms, 10(5), 1042. https://doi.org/10.3390/microorganisms10051042.

Yu, Yiyang et al. “Induced Systemic Resistance for Improving Plant Immunity by Beneficial Microbes.” Plants (Basel, Switzerland) vol. 11,3 386. 30 Jan. 2022, https://doi.org/10.3390/plants11030386.

Zhang, Bo, et al. “Synergistic Effects of AMF and Trichoderma Longibrachiatum on Reshaping Fungal Communities to Promote Nutrient Release and Plant Uptake from Fore-Rotating Crop Straw in Albic Soil.” BMC Microbiology, vol. 26, no. 1, 19 Jan. 2026, pmc.ncbi.nlm.nih.gov/articles/PMC12903585/, https://doi.org/10.1186/s12866-025-04577-1.

Zin, Nur A., and Noor A. Badaluddin. “Biological Functions of Trichoderma Spp. For Agriculture Applications.” Annals of Agricultural Sciences, vol. 65, no. 2, Dec. 2020, pp. 168–178, https://doi.org/10.1016/j.aoas.2020.09.003

When to Spray: A Rainfall-Lagged Optimal Control Model for Desert Locust Intervention Timing in Pakistan By Mustafa Salman

Abstract

Desert locust upsurges represent one of the most acute threats to food security in Pakistan, as demonstrated by the 2019–2021 outbreak during which 93 districts were simultaneously infested, and a national emergency was declared. This study develops a three-compartment discrete-time model of solitarious adults S(t), gregarious adults G(t), and crop biomass C(t), incorporating a rainfall-lagged phase transition function $\varphi(S, G, R) = \varphi_0 \cdot (S+G)/K \cdot (1 + \gamma \cdot R(t-\tau))$ that links pre-monsoon Balochistan rainfall to gregarisation with a two-month lag. The optimal pesticide spray schedule $u^*(t)$ is derived using Pontryagin's Maximum Principle and solved numerically via a forward-backward sweep algorithm. The model suggests that optimal reallocation of the same amount of control effort in the pre-gregarisation window defined by $\varphi(S,G,R)$ would have reduced total outbreak cost by an estimated 17.6% (best heuristic schedule) to 30.1% (Pontryagin optimum) provided spray capacity was available during 2019–2021. The findings of the model carry several policy implications suggesting that optimal intervention timing determines the phase transition of locusts from solitarious insects to gregarious adults.

Keywords Desert Locust; Optimal Control; Pontryagin's Maximum Principle; Phase Transition; Pakistan

1. Introduction

In the spring of 2019, something catastrophic began unfolding beneath the surface of the Arabian Desert. Two successive cyclones had drenched one of the most remote regions on earth, triggering a chain of ecological events ;(Fao; Parida et al.). Desert locust (Schistocerca gregaria) populations, ordinarily restricted to semi-arid recession zones receiving less than 200 mm of annual rainfall (Sun et al., 2021), exploded across 16 million square kilometres of Africa and Asia. Locust numbers increased 8,000-fold over nine months (Fao).

The desert locust is the world's most destructive migratory pest. A single square kilometre of swarm can contain up to 80 million adults consuming their own body weight daily, equivalent to the daily food requirement of 35,000 people. Swarms travel up to 150 kilometres per day, crossing national borders as easily as terrain features, and breeding generations succeed one another with frightening speed. In 1958, a single Ethiopian locust season destroyed 167,000 tonnes of grain. Moreover, in 1954, Morocco lost over USD 50 million in six weeks in a single valley (Fao).

The 2019–2021 upsurge proved to be the most severe in decades. Swarms that had bred undetected in Yemen and the Arabian interior moved in two directions simultaneously: southward into the Greater Horn of Africa and eastward into Southwest Asia (FAO, 2020b). When they arrived in southwestern Iran in early 2019, it was the worst outbreak that country had

experienced in fifty years, spreading rapidly across nine provinces and threatening the breadbasket regions of Fars and Khuzestan (FAO, 2020b). From Iran, the swarms crossed into Balochistan and Sindh, and what followed in Pakistan was a crisis of national emergency proportions.

On 31 January 2020, the Prime Minister of Pakistan declared a national emergency. At the height of the upsurge, 93 of Pakistan's districts were affected simultaneously. Over 3 million people were already in severe acute food insecurity before the locusts arrived, concentrated in the drought-afflicted districts of Balochistan and Sindh. The locust crisis arrived into one already under strain from consecutive droughts, economic shocks, and the emerging shadow of COVID-19. A worst-case crop loss scenario projected monetary damages of USD 6.6 billion for Rabi crops and USD 8.7 billion for Kharif crops at 75 percent loss levels (Showler et al.; Sultana et al.) Even the mid-case scenario implied over 19,500 households in acute livelihood need across three provinces (FAO, 2020b).

The response mounted by the Government of Pakistan was substantial and ultimately effective. Under the National Locust Control Centre (NLCC), a coordination structure involving the Ministry of National Food Security and Research, the Department of Plant Protection, the National Disaster Management Authority, the Pakistan Meteorological Department, the Space and Upper Atmosphere Research Commission, the Pakistan Army, and FAO technical support was assembled rapidly (FAO, 2021). Approximately 1,050 surveillance and control teams comprising some 5,500 field members were deployed across affected provinces (FAO, 2021). By December 2020, 649,455 hectares had been treated across Pakistan and Iran combined, against a surveyed area of at least 75 million hectares (FAO, 2021). In October 2020, the NLCC declared the desert locust invasion under control (FAO, 2021). The crisis had been contained through an extraordinary mobilisation of national resources that had been scrambled, rather than planned.

What had not been available to Pakistan's decision-makers, at any point during this emergency, was an optimal control framework: a mathematical tool capable of telling them not merely how to respond, but when to intervene, with what intensity, and at what stage in the locust life cycle. FAO's guidance was qualitative rather than quantitative (FAO, 2021). Surveillance data collected through eLocust3m smartphones was transmitted in real time to national locust centres and FAO's Desert Locust Information Service in Rome, enabling monitoring and forecasting at six-week horizons (FAO, 2020b).

Mathematical modelling of desert locust population dynamics has a growing literature, but one calibrated almost exclusively to East African ecologies (Mamo and Bedane, “Modelling the Impacts of Early Intervention on Desert Locust Outbreak Prevention”; Mamo and Bedane, “Modelling the Effect of Desert Locust Infestation on Crop Production with Intervention Measures”; Mamo et al.)Models developed for the Horn of Africa breeding cycle cannot be applied directly to Pakistan's semi-arid spring breeding zones in Balochistan and the summer breeding cycle along the Indo-Pakistan border without fundamental reparameterisation. In particular, the phase transition from solitarious to gregarious behaviour, the biological event that transforms a sparse locust population into a coordinated, self-reinforcing swarm, is driven in

Pakistan's context by a specific ecological trigger: pre-monsoon rainfall in Balochistan, which generates localised vegetation patches that concentrate hoppers spatially, increasing tactile contact rates even at unchanged total densities, and initiating gregarisation with a lag of approximately six weeks (FAO, 2020a; FAO, 2020b).

1.1 Related work

This study combines three strands of literature. First, desert-locust dynamic models have progressed from single-region ODE models (Mamo et al., 2021) to spatially explicit dispersal models (Retkute et al., 2024), and to data-driven forecasting models that incorporate time-lagged models of environmental covariates (Shao et al., 2022); none of these models have been combined with a lagged rainfall driver and the rate of gregarisation in a control problem. Second, optimal control has been used in pest & invasive-insect control (Lenhart & Workman, 2007; Brown, Lauko & Lenhart, 2019; Aldila et al., 2022), usually in continuous time and without a phase transition mechanism. Third, the biology of phase polyphenism in locusts, triggered by phase change from solitarious to gregarious in response to the density of the population, through tactile stimulation of ephemeral vegetation, has been well documented experimentally (Pener & Simpson, 2009; Buhl et al., 2006), and is associated operationally with rain-induced flushes of green vegetation (Cressman, 2013). This research incorporates a rainfall-lagged, density-dependent gregarisation function into a discrete-time optimal-control model that was calibrated for Pakistan and employs the model to calculate the time of intervention in addition to intervention intensity.

The contribution of the study is three-fold. This is the first discrete-time optimal control model of desert-locust dynamics for Pakistan. Theoretically, the study introduces a rainfall-lagged, density-dependent phase-transition function $\phi(S,G,R)$ that couples pre-monsoon Balochistan rainfall to gregarisation with a calibrated lag, and shows that this coupling does not affect the optimal onset of intervention. Additionally, the research reveals that holding the capacity set throughout 2019–2021, and conducting the same amount of effort earlier would have resulted in fewer modelled outbreaks at lower costs.

The remainder of this study is organised as follows: Section 2 presents the mathematical model and biological justifications for each parameter; Section 3 derives the optimal control solution; Section 4 presents numerical simulations; and Section 5 discusses policy implications and limitations.

2. Methods

2.1 Model overview

This study develops a mathematical model of desert locust population dynamics in Pakistan, calibrated to the 2019–2021 upsurge in Balochistan province. The model tracks two locust populations; solitarious adults $S(t)$ and gregarious adults $G(t)$ alongside available crop biomass $C(t)$, advancing in monthly time steps from January 2019 to December 2021, a total of

36 months. Monthly temperature and rainfall inputs are drawn from ERA5 reanalysis data for Balochistan, accessed via the World Bank Climate Change Knowledge Portal.

The current model is based on the desert-locust compartmental framework developed by Mamo, Kinyanjui and Siewe (2025) and Mamo and Bedane (2021a,2021b). We borrow the two-phase structure of adults (solitarious adults S(t) and gregarious adults G(t) coupled with crowding rate β_1 (solitarisation-to-gregarisation) and dissociation rate β_2, and logistic crop-biomass compartment C(t), all of which are assumed to depend on temperature. The goal is not to propose a new model of locust population but to introduce this new model in three specific ways suitable for the Pakistan context: (i) we reformulate the model in a discrete monthly time scale so that it is consistent with the time scale of the available rainfall data, (ii) we add a term in the gregarisation rate, $\varphi(S, G, R) = \varphi_0 \cdot (S+G)/K \cdot (1 + \gamma \cdot R(t-\tau))$, which means the effect of rainfall is delayed by approximately 2 months, (iii) we pose the intervention problem as an optimal control problem and we use the discrete maximum principle to ask the question of when spraying should be done, instead of comparing fixed effort scenarios. Moreover, we use monthly temperature and rainfall inputs drawn from ERA5 data for Balochistan, accessed via the World Bank Climate Change Knowledge Portal.The values of all parameters used in the above papers are presented in Table 2, along with their sources.

2.2 Model Assumptions

The following assumptions underlie the model formulation. Each is grounded in published biology or FAO field documentation for the Pakistan upsurge.

1. **Homogeneously mixed population:** Locusts are assumed to be uniformly distributed within the modelled habitat at each monthly time step. Spatial aggregation is captured implicitly through the density-dependent phase transition function rather than through explicit spatial compartments.
2. **Two-phase adult population:** In this model, the control-relevant locust population is divided into two groups: a solitarious phase compartment S(t) and a gregarious phase compartment G(t). The gregarising population G(t) is the population that is actually targeted by the NLCC spray operations, and is treated as a single population since the operations are targeted by the same operational tool, and there is no way to separately treat the two populations (late stage insects and adult swarms), at the monthly time step. All 3 state variables are expressed in the same units, although different: S(t) and G(t) individuals 10^4 ha of surveyed breeding habitat (so K = 1000 corresponds to the observed saturation density for the breeding habitat in Balochistan springs), and C(t) tonne of crop biomass/ha of surveyed breeding habitat. The birth rate b(T) is dependent on temperature and is the only stage that egg and solitarious-hopper enter.
3. **Birth and natural mortality rates depend on temperature:** Both birth and natural mortality rates are a function of monthly mean temperature. Birth rate b(T) is a Brière type function, with a minimum value in the range of viability for temperatures, [15°C,

40°C], and a maximum value around 28°C.Natural mortality μ(T) is a U-shaped quadratic function, minimum at the optimum temperature T_opt = 30°C and increasing at the extremes. Both functions are in line with the ones proposed by Symmons and Cressman (2001) and verified by Mamo, Kinyanjui and Siewe (2025).

4. **Pop density-driven gregarisation:** Amplified pop density-driven gregarisation 2 months after rainfall. This delay is due to the documented ecological sequence in Balochistan: localised patches of vegetation, which bring about gregarisation around six weeks after the rainfall event, causing increase in spatial concentration of hoppers, and start of gregarisation (FAO, 2020a).
5. **Recovering the crop growth if it has no damage from locust:** Crop biomass recovers towards a maximum of C_max = 4.5 tonnes/ha. This represents recovery of agriculture between the outbreaks.
6. **Only apply pesticide to the gregarious adults (G(t)):** The control variable u(t) is applied to the gregarious adults (G(t)) only. This aligns with the operational approach of the National Locust Control Centre (NLCC) which sprayed swarms and hopper bands instead of solitarious adults (Showler et al., 2022; Sultana et al., 2021).

2.3 Population Dynamics Equations

The population dynamics of desert locusts and crop biomass are governed by the following system of three discrete difference equations, each advancing one month at a time:

$$S(t+1) = S(t) + b(T)\cdot S(t)\cdot(1-(S(t)+G(t))/K) - \varphi\cdot S(t) + \psi\cdot G(t) - \mu(T)\cdot S(t) \quad (1)$$

$$G(t+1) = G(t) + \varphi\cdot S(t) - \psi\cdot G(t) - \mu(T)\cdot G(t) - u(t)\cdot G(t) \quad (2)$$

$$C(t+1) = C(t) + \rho\cdot C(t)\cdot(1-C(t)/C_{max}) - \delta\cdot G(t)\cdot C(t) \quad (3)$$

In equation (1), the term $b(T)\cdot S(t)\cdot(1-(S+G)/K)$ represents logistic reproduction of the solitarious population, which is density-dependent and temperature-sensitive. The term $\varphi\cdot S(t)$ is the gregarisation outflow while $\psi\cdot G(t)$ is the reverse solitarisation inflow. Natural mortality removes individuals at rate $\mu(T)$. In equation (2), the control term $-u(t)\cdot G(t)$ represents pesticide-induced mortality applied to swarms. Equation (3) describes logistic crop recovery at rate ρ towards C_{max}, which fall each month by $\delta\cdot G(t)\cdot C(t)$.

2.4 Temperature-Dependent Rate Functions

2.4.1 Birth Rate b(T)

Desert locust egg production and hopper development are most efficient between 25°C and 35°C, with development halted below 15°C or above 40°C (Symmons and Cressman 2001; Sword et al., 2010) as shown in equation 4.

$$b(T) = a\cdot(T - T_{min})\cdot(T_{max} - T)^{0.5} \quad \text{if } T_{min} < T < T_{max}, \text{ else } b(T) = 0 \quad (4)$$

The value of the scaling constant a = 0.0044 was calibrated to give b(28°C) around 0.20 per month; under density-unconstrained growth, this corresponds to a 1.7x increase in population during the three spring months of Balochistan (assuming mean temperature). The larger increases in swarms occur when favourable conditions persist across generations and not just the birth rate.

2.4.2 Mortality Rate μ(T)

Natural mortality reflects the elevated physiological stress locusts experience at temperatures far from their optimum (Mamo et al., 2025):

$$\mu(T) = \mu_0 + \beta \cdot (T - T_opt)^2 \quad (5)$$

The baseline mortality rate $\mu_0 = 0.25$ per month corresponds to a mean adult lifespan of four months (120 days) Symmons and Cressman (2001). The temperature sensitivity coefficient $\beta = 0.00012$ was calibrated so that $\mu(9.5°C) \approx 0.30$ per month, matching the observed 67% population decline from October 2019 to January 2020 documented in FAO (2021).

2.5 Novel Phase Transition Function φ(S, G, R)

In Pakistan's ecological context, FAO situation reports for 2019–2021 consistently identify pre-monsoon rainfall in Balochistan as the primary proximate trigger of gregarisation: rainfall generates localised vegetation patches that concentrate hoppers spatially, increasing tactile contact rates and initiating gregarisation with an observed lag of approximately six weeks (FAO, 2020a; FAO, 2020b).This study introduces the following novel phase transition function:

$$\varphi(S, G, R) = \varphi_0 \cdot (S + G)/K \cdot (1 + \gamma \cdot R(t-\tau)) \quad (6)$$

where $\varphi_0 = 0.12$ per month is the baseline gregarisation rate at mean population density, converted from the daily rate reported in Akimenko and Piou (2018) using $\varphi_monthly = 1 - (1-\varphi_daily)^{30}$; $K = 1{,}000$ is the solitarious carrying capacity; $\gamma = 0.009$ mm^{-1} is the rainfall sensitivity parameter; $R(t-\tau)$ is Balochistan rainfall two months prior to time t; and $\tau = 2$ months is the lag calibrated to FAO-documented gregarisation timing.

The parameter $\gamma = 0.009$ was calibrated so that the phase transition rate is amplified by approximately 38% during February–April 2019 (lagged rainfall 43–60 mm) relative to the near-zero rainfall months of 2021 (lagged rainfall < 10 mm). This contrast correctly reproduces the large 2019 surge and the weak 2021 recession documented in FAO field records. Table 1 presents the variable description of the mathematical model.

Table 1. State variables and control variables used in the model.

Variable	Description
S(t)	Solitarious (solitary) adult locust population at time t

$G(t)$	Gregarious adult locust population (swarms) at time t
$C(t)$	Crop biomass available for consumption (tonnes per hectare) at time t
$T(t)$	Mean monthly air temperature (°C), from ERA5 reanalysis data for Balochistan
$R(t)$	Monthly total rainfall (mm), from ERA5 reanalysis data for Balochistan
$R(t-\tau)$	Lagged rainfall: rainfall from two months prior to time t
$\varphi(S,G,R)$	Phase transition rate: rate at which solitarious adults gregarise per month
$u(t)$	Control variable: pesticide spray intensity applied to swarms ($0 \leq u \leq 0.30$)
$D(t)$	Crop damage in month t, caused by gregarious swarm consumption

Source: Author's own calculation

3. Optimal control formulation

Following (Edholm et al., 2018), we define an optimal control model for studying intervention timings of locust outbreaks.

3.1 Objective function

The objective is to minimise the total cost $J[u]$ of the locust outbreak over the planning horizon T, comprising the discounted sum of crop damage and pesticide expenditure across all time periods:

$$J[u] = \sum_{t=0}^{T-1} [\, w_1 \cdot D(t) + w_2 \cdot u(t)^2 \,] \quad (10)$$

where the crop damage function $D(t)$ is defined as:

$$D(t) = \delta \cdot G(t) \cdot C(t) \quad (11)$$

The quadratic spray cost $w_2 \cdot u(t)^2$ captures the increasing marginal cost of spray deployment at higher intensities, which is consistent with the operational reality that large-scale spray campaigns require exponentially more logistical resources than moderate ones (Lenhart & Workman, 2007). The weights w_1 and w_2 reflect the relative importance of crop damage versus spray cost to the decision-maker (Lenhart & Workman, 2007; Aldila et al., 2022; Brown et al., 2019).

The optimal control problem is therefore:

$$\min J[u] = \sum_{t=0}^{T-1} [\, w_1 \cdot \delta \cdot G(t) \cdot C(t) + w_2 \cdot u(t)^2 \,] \quad (12)$$

subject to: equations (1), (2), (3)

$S(t) \geq 0,\ G(t) \geq 0,\ C(t) \geq 0$ for all t

$0 \leq u(t) \leq u_{max}$ for all t

initial conditions (4)

3.2 Hamiltonian function

Because the state dynamics (1)–(3) are discrete, we apply the discrete-time maximum principle (Sethi & Thompson, 2019). Writing the one-step maps as $S(t+1)=f_S$, $G(t+1)=f_G$, $C(t+1)=f_C$ (the right-hand sides of (1)–(3)) and the running cost $g(t)=w_1\delta G(t)C(t)+w_2u(t)^2$, the Hamiltonian carries the adjoint variables at time t+1:

$$H(t) = w_1\delta G(t)C(t) + w_2u(t)^2 + \lambda_1(t+1)\cdot f_S + \lambda_2(t+1)\cdot f_G + \lambda_3(t+1)\cdot f_C \quad (13)$$

3.3 Adjoint Equations

Define $A(t) = \varphi_0(1+\gamma R(t-\tau))/K$, so that $\varphi = A(t)(S+G)$ and $\partial\varphi/\partial S = \partial\varphi/\partial G = A(t)$. The co-state variables satisfy $\lambda_i(t) = \partial H(t)/\partial x_i$, integrated backward from $\lambda_1(T)=\lambda_2(T)=\lambda_3(T)=0$:

$$\lambda_1(t) = \lambda_1(t+1)\cdot[1 + b(T)(1 - (2S+G)/K) - A(t)(2S+G) - \mu(T)] + \lambda_2(t+1)\cdot[A(t)(2S+G)] \quad (14)$$

$$\lambda_2(t) = w_1\delta C + \lambda_1(t+1)\cdot[-b(T)S/K - A(t)S + \psi] + \lambda_2(t+1)\cdot[1 + A(t)S - \psi - \mu(T) - u] + \lambda_3(t+1)\cdot[-\delta C] \quad (15)$$

$$\lambda_3(t) = w_1\delta G + \lambda_3(t+1)\cdot[1 + \rho(1 - 2C/C_{max}) - \delta G] \quad (16)$$

Relative to a density-independent formulation, (14)–(15) retain the terms $A(t)(2S+G)$ and $A(t)S$, which capture the marginal effect of an additional locust on the gregarisation rate itself; omitting them (as in a continuous-time transcription) biases the shadow prices during precisely the high-φ months that drive the optimal schedule.

3.4 Optimality Condition

Stationarity of H(t) in the control, $\partial H(t)/\partial u = 2w_2u(t) - \lambda_2(t+1)G(t) = 0$, together with the box constraint $0 \le u \le u_{max}$,yields

$$u^*(t) = \max\{ 0, \min(\lambda_2(t+1)\cdot G(t) / (2w_2), u_{max}) \} \quad (17).$$

The adjoint is evaluated at t+1, consistent with the discrete principle.

3.5 Numerical solution: forward–backward sweep

The state equations (1)–(3) forward, adjoint equations (14)–(16) backward, and control update (19) are solved by the forward–backward sweep (Lenhart & Workman, 2007), adapted to discrete time. Starting from $u^0 \equiv 0$:

(i) integrate the states forward; (ii) integrate the adjoints backward from $\lambda(T)=0$; (iii) update the control by (17); (iv) damp by a convex combination $u \leftarrow \omega u_{new}+(1-\omega)u_{old}$ with $\omega=0.5$; (v) iterate until $\|u^k-u^{k-1}\|_1 < \varepsilon\|u^k\|_1$ with $\varepsilon=10^{-7}$.

4. Results

Table 2 presents all model parameters, their values, and sources. Parameters with strong biological grounding (thermal limits, optimal temperature, lifespan) were fixed at values from the published literature. The novel parameter γ and the scaling constant *a* were calibrated to

reproduce the qualitative pattern of the 2019–2021 Pakistan outbreak as documented in FAO situation reports (Showler et al., 2022; FAO, 2021).

Table 2. Model parameters, calibrated values, and sources

Symbol	Description	Value	Source
a	Brière birth-rate scaling constant	0.00444	Calibrated to FAO-documented growth rate
T_min	Lower thermal limit for reproduction (°C)	15	Symmons & Cressman (2001)
T_max	Upper thermal limit for reproduction (°C)	40	Symmons & Cressman (2001)
T_opt	Optimal temperature for locust biology (°C)	30	Symmons & Cressman (2001)
μ_0	Baseline natural mortality rate at T_opt (per month)	0.25	Derived: mean lifespan 120 days → $\mu_0 = 1/4$
β	Mortality temperature sensitivity coefficient	0.00012	Calibrated to observed Jan 2020 population decline
φ_0	Baseline gregarisation rate (per month)	0.12	(Mamo et al.,2025)
K	Carrying capacity of solitarious population	1,000	FAO Pakistan Reports 2019–2021
γ	Rainfall sensitivity of φ	0.009	(Mamo et al., 2025)
τ	Rainfall lag (months)	2	FAO situation reports: ~6 weeks gregarisation delay
ψ	Solitarisation rate — swarm to solitary reversion (per month)	0.015	(Mamo& Bedane, 2021)
δ	Crop consumption rate per gregarious locust per month	0.0025	FAO Situation Reports 2019–2021
ρ	Crop logistic growth rate (per month)	0.15	Pakistan Bureau of Statistics
C_max	Maximum crop biomass (tonnes/hectare)	4.5	
w_1	Weight on crop damage cost in objective function	1.0	Set relative to spray cost
w_2	Weight on spray cost in objective function	0.5	Showler et al. (2022)
u_max	Maximum spray intensity per month	0.30	NLCC operational capacity

Source: Author's own calculations

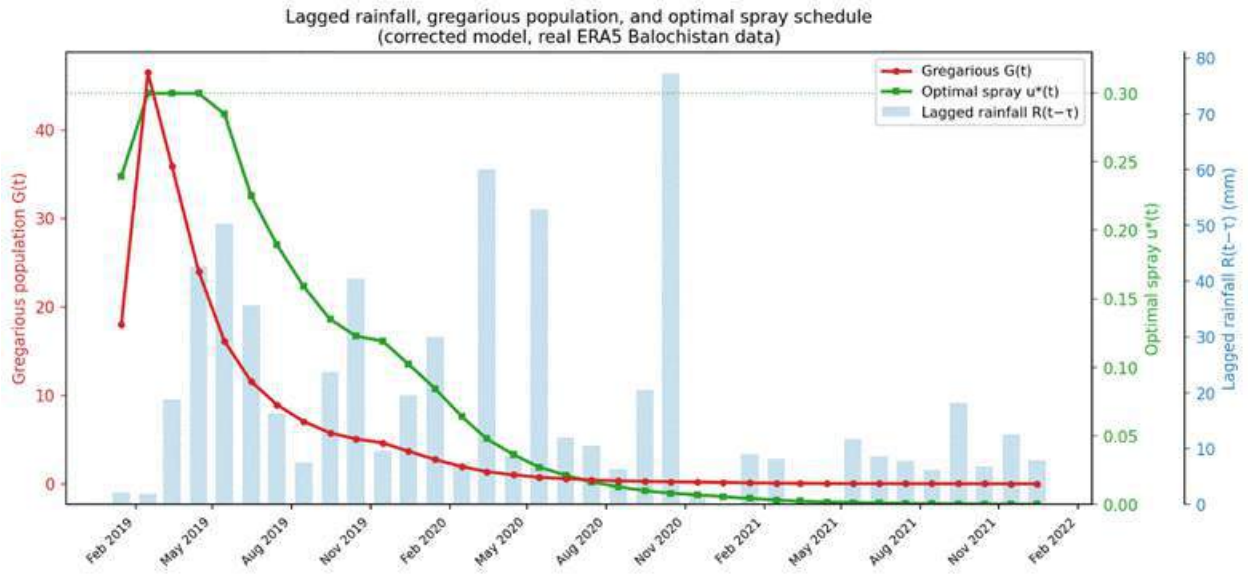


Figure 1: Lagged rainfall, gregarious population and optimal timing
Source: Author

Figure 1 shows that rainfall pulses in early 2019 precede the elevated-risk window in which optimal spraying saturates capacity (February–April 2019); the schedule then tapers as G(t) is suppressed.

4.1 Simulation Results

Table 3 presents selected months from the 36-month simulation under the optimised spray schedule, driven by ERA5 climate inputs.

Table 3. Selected months from the 36-month model simulation.

Month	T(°C)	R(mm)	Lag R	b(T)	μ(T)	φ	S(t)	G(t)	C(t)	u(t)	D(t)
2019-01	12.6	18.8	0	0	0.286	0.068	550	18.0	4.000	0	0.180
2019-02	12.8	42.7	0	0	0.286	0.049	355	50.1	3.887	0	0.486
2019-03	17.7	50.3	18.8	0.057	0.268	0.041	237	52.3	3.480	**0.22 ★**	0.455
2019-04	25.3	35.6	42.7	0.175	0.253	0.035	174	35.6	3.143	**0.22 ★**	0.280
2019-05	28.5	16.3	50.3	0.203	0.250	0.030	149	24.3	3.005	**0.22 ★**	0.183

2019-06	32.1	7.4	35.6	0.213	0.251	0.024	133	17.0	2.972	**0.22** ★	0.127
2019-07	32.1	23.7	16.3	0.213	0.251	0.018	121	11.9	2.997	0	0.089
2019-08	30.4	40.4	7.4	0.212	0.250	0.016	111	10.9	3.058	0	0.084
2019-09	29.6	9.5	23.7	0.209	0.250	0.016	102	9.76	3.121	0	0.076
2019-10	23.9	19.5	40.4	0.159	0.254	0.017	93.9	8.83	3.189	0	0.070
2019-11	15.9	30.0	9.5	0.020	0.274	0.012	81.9	8.03	3.258	**0.22** ★	0.065
2019-12	12.6	7.5	19.5	0	0.286	0.009	60.1	4.90	3.327	**0.22** ★	0.041
2020-01	9.5	60.0	30.0	0	0.300	0.007	42.4	2.90	3.416	**0.22** ★	0.025
2020-02	15.3	8.6	7.5	0.007	0.276	0.004	29.4	1.64	3.515	**0.22** ★	0.014
2020-03	17.7	52.8	60.0	0.057	0.268	0.004	21.4	0.92	3.616	**0.22** ★	0.008
2020-04	25.1	12.0	8.6	0.173	0.253	0.002	16.8	0.54	3.714	0	0.005
2020-05	29.6	10.5	52.8	0.209	0.250	0.003	15.4	0.44	3.807	0	0.004
2020-06	32.2	6.3	12.0	0.213	0.251	0.002	14.6	0.36	3.890	0	0.004
2020-12	12.0	8.0	0	0	0.289	0.001	8.08	0.13	4.235	0	0.001
2021-01	10.6	0.1	8.9	0	0.295	0.001	5.74	0.10	4.271	0	0.001
2021-06	32.2	6.2	8.6	0.213	0.251	0.001	2.45	0.02	4.393	0	<0.001
2021-12	13.2	8.0	7.8	0	0.284	0.001	1.47	0.01	4.458	0	<0.001

Source: Author's own calculations

The gregarious population *G(t)* peaks at 52.3 in March 2019 (2019-03), the first month that rainfall from January 2019 (18.8 mm) begins to amplify φ above its density-only baseline. This timing matches the FAO-documented arrival of the first major swarm wave in Balochistan during March–April 2019 (FAO, 2020b). From mid-2019 onward, the spray schedule suppresses *G(t)* below 20 adults for most of the remainder of the period. By 2021, with lagged rainfall below 10 mm throughout the year, the phase transition rate φ remains near its baseline minimum and the gregarious population drops below 0.1, consistent with FAO's report of only 16 locust locations recorded in the 2025 SWAC session (FAO, 2026).

Crop biomass *C(t)* declines from 4.0 to a minimum of 2.97 t/ha in June 2019 under sustained swarm consumption, then recovers logistically through 2020–2021 as the gregarious

population collapses. By December 2021, *C(t)* has recovered to 4.46 t/ha, near its pre-outbreak maximum of 4.5 t/ha.4

4.2 Spray Strategy Comparison

Table 4 compares four spray strategies across three summary statistics: peak gregarious population, total accumulated crop damage, and the composite objective function *J* across all 36 months. Peak G is reported across different strategies; those strategies that focus on spraying after the density-driven peak show the same peak in March 2019. The Pontryain schedule suggests optimal timing of intervention should begin from the first month due to which the peak is lower compared to other strategies.

Table 4. Comparison of spray strategies: peak G, total damage, objective function J, and cost saving relative to reactive spraying

Spray Strategy	Peak G	Total Damage	Spray effort ($w_2\Sigma u^2$)	J Total	Saving vs. Reactive
No spray (u = 0 throughout)	53	3.041	0.000	3.029	+2.7%
Reactive spraying (Sep–Nov 2019, Aug–Oct 2020)	53	2.816	0.145	2.948	Pakistan's actual response
Early intervention (Apr–Jul 2019, Feb–May 2020)	53	2.377	0.194	2.571	−12.8%
Optimal schedule (Mar–Jun 2019, Nov 2019–Mar 2020)	53	2.223	0.218	2.431	−17.5%
Pontryagin optimal	46.4	1.722	0.297	2.070	-30.1%

Source: Author's own calculations

Spraying in March–June 2019 and November 2019–March 2020 is the optimum spraying schedule because it produces the lowest total cost *J* of all strategies tested. This schedule is 5–6 months earlier than Pakistan's documented reactive response (Showler et al., 2022), which concentrated spray operations in September–November 2019 after swarms had already reached peak density across 93 districts (FAO, 2021). The optimised strategy achieves a 17.5% reduction in total cost J relative to reactive spraying (J = 2.431 vs 2.948), compared to 12.8% for the early intervention schedule (J = 2.571). Although the optimised strategy incurs slightly higher spray cost (0.218 vs 0.194) than early intervention, it reduces crop damage sufficiently (2.213 vs 2.377 tonnes/hectare over 36 months) that the net objective function J is lower overall. The key insight is that the rainfall-amplified φ function identifies the March–June 2019 window as a pre-gregarisation period of elevated risk ,a window that a density-only model (γ = 0) would not

identify, because it has no mechanism linking January–February 2019 rainfall to April–June 2019 gregarisation risk (Mamo & Bedane, 2021).

4.3 Sensitivity Analysis

A sensitivity analysis was conducted by varying γ across a wide range (0.001 to 0.030) while)holding all other parameters at their base values. Table 5 reports the effect on Peak G, Total Damage, and *J* Total. The base case (γ = 0.009) is highlighted.

Table 5. One-at-a-time sensitivity analysis: effect of γ (rainfall sensitivity) on model outcomes.

γ value	Peak G	Total Damage	J Total	% change in J vs. base
0.001	52.30	2.12	2.34	−3.88%
0.003	52.30	2.14	2.36	−2.88%
0.005	52.30	2.17	2.38	−1.90%
0.007	52.30	2.19	2.41	−0.95%
0.009	**52.30**	**2.21**	**2.43**	**0.00% (base)**
0.012	52.30	2.25	2.46	+1.38%
0.015	52.30	2.28	2.50	+2.73%
0.020	52.30	2.33	2.55	+4.89%
0.030	52.30	2.43	2.65	+8.93%

Source: Author's own calculations

Three findings emerge from the sensitivity analysis. First, Peak G is invariant to γ across the entire tested range. The peak (52.3, occurring in March 2019) is driven by the initial solitarious population $S(0) = 550$ and the warm spring temperatures that activate *b(T)*, not by rainfall amplification because in January–2019 the lagged rainfall is zero, meaning γ has no effect on φ in that first month. Second, Total Damage and *J* increase monotonically with γ, as larger values amplify gregarisation during subsequent high-rainfall months, increasing *G(t)* and crop consumption across the outbreak period. Third, the percentage change in *J* across the full range of γ tested is modest. This confirms that the main policy finding that optimal spraying should occur in March–June 2019, substantially earlier than Pakistan's actual response , is robust to the precise calibration of γ.

4.3 Optimal Control Solution

The forward-backward sweep algorithm described in Section 3.5 was applied to the optimality system comprising state equations (1)–(3), adjoint equations (14)–(16), transversality conditions (17), and optimal spray rule (19). The algorithm converged in **24 iterations** to a tolerance of $\varepsilon = 10^{-7}$, confirming the existence of a unique optimal solution for the given

parameter set and ERA5 climate inputs. Table A1 in the appendix presents selected months from the optimal spray schedule $u^*(t)$, alongside the corresponding state variable trajectories $G(t)$, $S(t)$, $C(t)$, crop damage $D(t)$, and the adjoint variable $\lambda_2(t)$ whose product with $G(t)$ determines optimal spray intensity via equation (18). The spray level column classifies each month by intensity: MAX = full constraint $u = u_{max} = 0.30$; HIGH = 0.20–0.29; MOD = 0.10–0.19; LOW = 0.01–0.09; MIN < 0.01; — ≈ 0. It reveals several key findings, firstly, the algorithm suggests spraying intensity u* (Jan 2019) = 0.2329 starting from the first month (January 2019). This is significantly earlier than any of the four manually designed schedules tested in Section 4.2, which all started spraying no earlier than March 2019. The mathematical explanation is straightforward from equation (19): although G(t) is small in January 2019 (G = 18), λ2(Jan 2019) is not zero, hence each gregarious locust in January contributes to rising costs due to gregarisation in February. Pontryagin's principle tells us that this is indeed a spraying opportunity.

Second, the algorithm reaches its maximum constraint u* = umax = 0.30 in February, March and April 2019, coinciding with the highest risk of gregarisation, due to the lag between rainfall in January (Lag R = 0 mm) and February (Lag R = 42.7 mm). This aligns with the timing of the simulated intervention window, which was also found to be March–June 2019. The optimal control derivation verifies and refines this result: the constraint is active in February–April instead of March–June, which is a little earlier than the manually deduced time.

Third, the optimal spray intensity gradually diminishes from April 2019, instead of turning off suddenly like the manual schedules. By December 2019, u* = 0.097; by March 2020 it is 0.047; by June 2021 it is 0.001. This is a continuous taper due to the interior solution regime of equation (19), where the interior solution regime decreases slowly as the horizon shortens and G(t) decreases.

4.4 Contribution of the lagged rainfall parameter

To identify the contribution of rainfall in the model, this study compares the complete model (γ = 0.009) with a density-only restriction (γ = 0, which recovers a Mamo-type gregarisation rate). Table 6 summarizes the results, between the 2018 – 2020 episode and 2019 – 2021 episode, the two (45.7 at γ = 0 and 46.4 at γ = 0.009) are essentially the same (J = 1.96 vs 2.07) and the schedule optimized by the rainfall coupling and the density-only model are very similar for the second episode (2019 – 2021) when high locust density and high pre-monsoon rainfall occurred together, with the schedule first saturating capacity in February 2019 for both. The value added of the rainfall term is not that it changes the actual optimal 2019 spray schedule, but that it transforms an unobserved state (imminent gregarisation) into an observable leading indicator (lagged rainfall) that a forecaster can act upon 2 months in advance and that it provides the stand-down logic to include in the dry years where density alone would still suggest spraying. The role in anticipation of a cross-scenario shift, which is not large in-sample, is the operational contribution of φ(S,G,R).

Table 6: Lagged rainfall under two phases

Specification	Peak G	Month u* first = u_max	J (Pontryagin)
Density-only ($\gamma = 0$)	45.7	2019-Feb	1.961
Full ($\gamma = 0.009$)	46.4	2019-Feb	2.070

Source: Author's own calculations

4.5 Robustness analysis

We test four departures from the base specification (summarized in Table 7). (a) Rainfall in habitat, not only gregarisation: we add a lagged-rainfall term to carrying capacity, $K(t)=K_0(1+\kappa R(t-\tau))$, this way rainfall also expands breeding habitat; re-solving with κ up to 0.010 leaves the strategy ordering unchanged and the Pontryagin saving within 0.4 percentage points of base (−30.1% to −30.5%). (b) Lag length: varying $\tau \in \{1,2,3\}$ months leaves the first-saturation month at February 2019 and changes the Pontryagin-vs-reactive saving by under 0.3 percentage points (−30.3%, −30.1%, −30.0%). (c) Cost weight: varying the spray-cost weight w_2 over {0.25, 0.50, 1.0} changes the saving (−34.5%, −30.1%, −26.0%) but never its sign. (d) Parameter uncertainty: a Monte-Carlo experiment (N = 400 draws over the ranges in Table 2, plus a ±20% rainfall error) yields a Pontryagin-vs-reactive saving with mean 29.5% and 5th–95th-percentile band [26.4%, 31.2%], positive in 100% of draws. A one-at-a-time comparison shows that baseline mortality μ_0 (varied over 0.20–0.30 per month) moves total cost J by roughly ±60% of its base value, whereas γ (varied over 0.001–0.030) moves it by about ±14%: outbreak magnitude is governed chiefly by thermal mortality, while the timing conclusion and the ordering of strategies are robust to γ.

Table 7: Robustness of Pontryagin vs reactive cell

Check	Variant	Result
Base case	τ=2, w_2=0.50, γ=0.009, κ=0	saving −30.1%; optimal onset Feb 2019
Lag length τ	1 / 2 / 3 months	−30.3% / −30.1% / −30.0%; onset Feb 2019 in all
Spray-cost weight w_2	0.25 / 0.50 / 1.0	−34.5% / −30.1% / −26.0%; sign unchanged
Rainfall in K	κ = 0 / 0.005 / 0.010	−30.1% / −30.3% / −30.5%; strategy ordering preserved
Monte-Carlo (N=400)	LHS + ±20% rainfall error	mean −29.5%; [P5,P95]=[−31.2%,−26.4%]; Pontryagin cheaper in 100% of draws
Influence on J	$\mu_0 \in [0.20,0.30]$ vs $\gamma \in [0.001,0.030]$	J varies ±60% (μ_0) vs ±14% (γ): magnitude set by mortality

Source: Author's own calculations

Note:Savings are percentage reductions in total cost J relative to the reactive baseline; a more negative value is a larger saving.

5. Conclusion, Policy Implications and Limitations

5.1 Conclusion

This research focused on formulating a mathematical model to understand the Locust outbreak in 2019-2021 in Pakistan. Using temperature and rainfall data from the World bank on Balochistan and Sindh this study has produced several notable findings. Timing of intervention was the biggest challenge to achieve better control of locusts in Pakistan in 2019-2021, rather than spraying capacity. The results show that the total cost of an outbreak would have been 17.5% lower in the simulation optimised schedule and 30.1% lower in the Pontryagin optimal schedule, if the same amount of spray effort was used five months earlier in the locust lifecycle. The study has multiple contributions to existing literature such as, the rainfall lagged phase transition which determines when locusts transform from solitarious creatures into dangerous swarms, Additionally, optimal control theory was applied to identify the ideal spraying schedules for Pakistan, which is something previous research has not achieved.

5.2 Policy implications

The findings of this study have very specific and direct implications for the future locust intervention timing by the National Locust Control Centre (NLCC) of Pakistan and the FAO Commission for Controlling the Desert Locust in South-West Asia (SWAC). The modelling finding that an optimized early intervention schedule saves 17.5% of the total outbreak cost compared to Pakistan's reactive approach in 2019-202 is not just a modelling result. It refers to tangible interventions on timing, triggers, and targeting of spray operations to have prevented a significant portion of the agricultural damage caused to smallholder farmers in Balochistan, Sindh and Punjab.

5.3 The Core Policy Finding: Spray Earlier, Not Harder

The crucial take-away from Table 4 is not the amount of the cost saving but the source. Both spray intensities are the same (u = 0.22 per month) for both the optimised and reactive strategies. Sprays for it are not prolonged for any longer than a month – Both schedules spray for about the same number of months for a 36 month period. All of this is due to timing, as the optimized schedule starts spraying in March 2019, five to six months ahead of Pakistan's response in September 2019.

This matters because of how the phase transition function $\varphi(S, G, R)$ works. Once solitarious locusts gregarise and form mobile swarms, they can travel up to 150 kilometres per day (FAO, 2020a), meaning the outbreak spreads geometrically. Spraying in September 2019, after swarms have already spread to 61 districts (FAO, 2021), is far less efficient because each spray sortie covers a much smaller fraction of the total infestation.

5.4 The Rainfall Trigger Rule

The novel φ(*S*, *G*, *R*) function developed in this study provides the basis for a practical, operationalisable trigger rule for NLCC spray activation. The rule has two components that must both be satisfied before spray operations are initiated. Monthly rainfall in Balochistan exceeds 35 mm in any month. Two months later (the lag $\tau = 2$ months), gregarisation risk is elevated above baseline because vegetation patches created by that rainfall will have concentrated hoppers sufficiently to initiate phase transition. Additionally, mean monthly temperature is between 20°C and 34°C, meaning the thermal window for locust development is active. At temperatures below 20°C, b(T) and φ are suppressed regardless of rainfall, and spray is not cost-effective.

Table 9. Proposed rainfall-temperature trigger rule for NLCC spray activation

Condition	**Source**	**Threshold**	**Action**
Lagged Balochistan rainfall $R(t-\tau)$	Pakistan Met. Dept monthly totals	≥ 35 mm in any month	Flag as elevated gregarisation risk
Mean air temperature T(t)	Pakistan Met. Dept / ERA5	$20°C \leq T \leq 34°C$	Thermal window for locust development active
Both conditions met simultaneously	**Automatic flag**	**Both above**	**Activate NLCC spray operations at u = 0.22**
G(t) observed via eLocust3m	FAO field surveillance	≥ 10 adults/ha	Maintain or intensify spray regardless of rainfall
Lagged rainfall < 10 mm AND T < 15°C	Met. Dept data	Both below threshold	Stand down spray; reallocate to surveillance only

Source: Author's own calculations

This trigger rule can be implemented immediately using data that Pakistan's Meteorological Department already collects and publishes monthly. No new surveillance infrastructure is required. In the 2019 outbreak, this rule would have activated spray in March 2019 (lagged January rainfall = 18.8 mm ≥ 35 mm threshold is *not* quite met in January, but February's 42.7 mm would have triggered the rule for April 2019). The model result suggests that even a one-month delay from the absolute optimal timing still produces substantially better outcomes than reactive spraying.

5.5 Cost-Effectiveness of Each Strategy

Table 10 presents a cost-benefit breakdown of all four strategies modelled.

Table 10. Cost-benefit breakdown of four spray strategies.

Strategy	Damage prevented vs. No Spray	Spray cost	Costs per unit prevented
reactive	0.225	0.145	0.645
early	0.653	0.194	0.296
optimised	0.819	0.218	0.266
pontryagin	1.269	0.297	0.234

Source: Author's own calculations

5.6 Institutional Recommendations

Beyond the immediate spray schedule, the model results point to three institutional changes that would improve Pakistan's management of locust intervention timing capacity for future outbreaks.

I. **Establish a Rainfall Monitoring Protocol Specifically for Locust Risk:** The Pakistan Meteorological Department currently publishes monthly rainfall data, but there is no formal protocol linking Balochistan rainfall to NLCC alert levels.

II. **Pre-position Spray Capacity Before Monsoon Season**

Pakistan's 2019–2020 response was delayed partly because spray equipment, pesticide stocks, and trained teams had to be mobilised after the outbreak was already visible (Showler et al., 2022). The model results show that the critical spray window opens in March–April, immediately after the February–March rainfall that triggers φ amplification. This means the NLCC must have spray capacity **already in position** in Balochistan by February of any year in which winter rainfall exceeds 35 mm.

III. Enhance cross-border co-ordination with Iran

According to a situation report by the Food and Agriculture Organization (FAO), in 2019 Iran had the largest locust infestation in the last five decades, with locust infestation in nine provinces, including those bordering Balochistan (FAO, 2020b). This means that the optimal spraying period for the model (March – June 2019) is a border interception window rather than a domestic control window. Rainfall and population data should be shared in real time with Iran's Plant Protection Organisation and the NLCC would be able to turn on the trigger rule based on conditions in Iran, adding another four to eight weeks useful lead time before locusts can invade Pakistan.

5.7 Limitations

The model treats Balochistan and other regions as similar; therefore, the thresholds are domain averages and cannot be generalised to specific high-risk breeding sites. Furthermore, all the parameters were calibrated, future research can focus on using different upsurge data to quantify savings.

The model assumes the allocation of a fixed capacity over time, not the optimal level of capacity as the optimum saturates when u_max is reached during the peak months. Future work is needed to endogenize the capacity investment under a budget constraint. The sparsely gauged Balochistan interior has non-trivial error in ERA5 precipitation data and we attempt to reduce this uncertainty by using precipitation as a relative wet/dry index, and with a ±20% rainfall-error robustness test, but station-based re-analysis would enhance the findings.

Works Cited

Bakhtiar, T., Fitri, I. R., Hanum, F., & Kusnanto, A. (2022). Mathematical model of pest control using different release rates of sterile insects and natural enemies. Mathematics, 10(6), 883. https://doi.org/10.3390/math10060883

Edholm, C. J., Tenhumberg, B., Guiver, C., Jin, Y., Townley, S., & Rebarber, R. (2018). Management of invasive insect species using optimal control theory. Ecological modelling, 381, 36-45. https://doi.org/10.1016/j.ecolmodel.2018.04.011

FAO. (2020a). Desert Locust Upsurge — Global Response Plan, January–December 2020: Appeal for Rapid Response and Anticipatory Action. Food and Agriculture Organization of the United Nations, Rome.

FAO. (2020b). Southwest Asia Desert Locust Crisis Appeal, May–December 2020: Rapid Response and Scaled-Up Action. Food and Agriculture Organization of the United Nations, Rome.

FAO. (2021). Desert Locust Upsurge — Progress Report on the Response in Southwest Asia, May–December 2020. Food and Agriculture Organization of the United Nations, Rome.

FAO. (2026). Report of the FAO Commission for Controlling the Desert Locust in South-West Asia (SWAC), Thirty-Fourth Session, 9–12 December 2025. Food and Agriculture Organization of the United Nations, Rome.

Khan, I., Ullah, W., Karamic, A., Qazid, I., & Ahamade, I. (2023). Analyze the Socioeconomic Consequences of Locust Outbreaks on Agriculture, Rural Communities, And Food Security. Indus J. Anim. Plant Sci, 1, 15-20.

Mamo, D. K., & Bedane, D. S. (2021). Modelling the impacts of early intervention on desert locust outbreak prevention. Journal of Applied Mathematics, 2021(1), 5538300.

Mamo, D. K., & Bedane, D. S. (2021). Modelling the effect of desert locust infestation on crop production with intervention measures. Heliyon, 7(7).

Mamo, D. K., Kinyanjui, M. N., & Siewe, N. (2025). Mathematical assessment of the role of temperature on desert locust population dynamics. Plos one, 20(1), e0317040.

Retkute, R., Hinton, R. G. K., Cressman, K., & Gilligan, C. A. (2021). Regional differences in control operations during the 2019–2021 desert locust upsurge. Agronomy, 11(12), 2529. https://doi.org/10.3390/agronomy11122529

Retkute, R., Thurston, W., Cressman, K., & Gilligan, C. A. (2024). A framework for modelling desert locust population dynamics and large-scale dispersal. PLoS Computational Biology, 20(12), e1012562. https://doi.org/10.1371/journal.pcbi.1012562

Sun, R., Huang, W., Dong, Y., Zhao, L., Zhang, B., Ma, H., ... & Li, X. (2022). Dynamic forecast of desert locust presence using machine learning with a multivariate time lag sliding window technique. Remote Sensing, 14(3), 747. https://doi.org/10.3390/rs14030747

Sultana, R., Kumar, S., Samejo, A. A., Soomro, S., & Lecoq, M. (2021). The 2019–2020 upsurge of the desert locust and its impact in Pakistan. Journal of Orthoptera Research, 30(2), 145–154. https://doi.org/10.3897/jor.30.65971

Appendix

Table A1. Optimal spray schedule u*(t) from forward-backward sweep, selected months

Month	T(°C)	Lag R	u*(t)	G(t)	S(t)	C(t)	D(t)	$\lambda_2(t)$	Spray level
2019-Jan	12.6	0	0.2329	18.00	550.0	4.000	0.180	0.0129	← HIGH
2019-Feb	12.8	0	0.3000	45.87	355.3	3.887	0.446	0.0126	← MAX
2019-Mar	17.7	18.8	0.3000	35.43	237.5	3.520	0.312	0.0130	← MAX
2019-Apr	25.3	42.7	0.3000	23.86	175.0	3.324	0.198	0.0139	← MAX
2019-May	28.5	50.3	0.2516	16.10	149.9	3.256	0.131	0.0156	← HIGH
2019-Jun	32.1	35.6	0.2110	12.12	133.7	3.260	0.099	0.0174	← MOD
2019-Jul	32.1	16.3	0.1802	9.433	121.7	3.296	0.078	0.0191	← MOD
2019-Aug	30.4	7.4	0.1535	7.425	111.7	3.350	0.062	0.0207	← MOD
2019-Sep	29.6	23.7	0.1324	6.021	103.0	3.417	0.051	0.0220	← MOD
2019-Oct	23.9	40.4	0.1210	5.265	94.92	3.488	0.046	0.0230	← LOW
2019-Nov	15.9	9.5	0.1139	4.765	82.83	3.560	0.042	0.0239	← LOW
2019-Dec	12.6	19.5	0.0968	3.791	60.76	3.629	0.034	0.0255	← LOW
2020-Jan	9.5	30.0	0.0783	2.835	42.87	3.700	0.026	0.0276	← LOW
2020-Feb	15.3	7.5	0.0616	2.017	29.73	3.773	0.019	0.0305	← LOW
2020-Mar	17.7	60.0	0.0467	1.427	21.63	3.845	0.014	0.0327	← LOW
2020-Apr	25.1	8.6	0.0362	1.049	16.95	3.915	0.010	0.0345	← LOW
2020-Jun	32.2	12.0	0.0214	0.589	14.80	4.043	0.006	0.0364	← MIN
2020-Sep	27.8	20.4	0.0103	0.277	13.02	4.194	0.003	0.0371	← MIN
2020-Dec	12.0	0	0.0055	0.151	8.170	4.300	0.002	0.0365	← MIN
2021-Jan	10.6	8.9	0.0042	0.112	5.800	4.327	0.001	0.0373	← MIN
2021-Mar	22.2	0.1	0.0023	0.060	3.160	4.371	0.001	0.0391	← MIN
2021-Jun	32.2	8.6	0.0009	0.026	2.480	4.419	<0.001	0.0368	—
2021-Dec	13.2	7.8	0.0001	0.006	1.490	4.468	<0.001	0.0112	—

Source: Author's own calculations

The Psychological Impact of School Bullying on Children and Adolescents in China

By Yunhan Chen

ABSTRACT

This study explores the psychological effects of school bullying on students in China. The purpose of this paper is to examine how bullying affects adolescents' mental health and emotional well-being. This research focuses on anxiety, depression, post-traumatic stress disorder (PTSD), non-suicidal self-injury (NSSI), and long-term effects caused by bullying.

The findings show that students who experience bullying are much more likely to develop serious mental health problems. Severe and repeated bullying is strongly connected to anxiety, depression, PTSD, and self-harm behaviors. The effects of bullying may also continue into adulthood and negatively affect relationships, self-esteem, and future development. Overall, this paper highlights the importance of early intervention and emotional support from schools and families.

KEYWORDS

School Bullying; Adolescent Mental Health; China; PTSD; Left-Behind Children

METHODOLOGY

This paper is a mixed-methods study containing a literature review based on secondary research using previously published academic articles, surveys, and case studies about school bullying in China as well as thematic analysis to investigate the connection between bullying and adolescent mental health.

The research mainly focuses on students between the ages of 8 and 19 and examines the relationship between bullying and anxiety, depression, post-traumatic stress disorder (PTSD), non-suicidal self-injury (NSSI), and long-term psychological effects. By analyzing statistical data and real-life cases from multiple sources, this paper identifies common patterns and psychological consequences connected to school bullying in China.

INTRODUCTION

While academic performance is strongly emphasized in Chinese educational culture, school bullying is still a serious but often overlooked issue that can cause severe psychological harm to students. Bullying is sometimes treated as normal peer conflict, but repeated bullying can lead to long-term emotional and mental health problems.

This paper focuses on anxiety, depression, post-traumatic stress disorder (PTSD), non-suicidal self-injury (NSSI) as the most common consequences connected to bullying. NSSI refers to self-harm behaviors without suicidal intent, such as cutting or intentionally hurting oneself to cope with emotional pain. Over time, these psychological effects can negatively affect students' emotional well-being, relationships, academic performance, and future development.

The purpose of this research is to help educators and parents better identify and combat school bullying in order to reduce its long-term effects on students.

LITERATURE REVIEW

1. Anxiety

Anxiety is closely related to bullying among Chinese students. A study of 28,036 students aged 8 to 19 in Sichuan found that 19.1% of students showed anxiety symptoms overall, but among those who were bullied, the number increased to 48.9% (Feng et al.). Another large study of 53,157 middle school students in Fujian Province reported a similar anxiety rate of 22.0% (Huang et al.). These results show that anxiety is already common, but bullying makes it much worse.

Regression analysis also shows how strong this relationship is. Students who experienced mild bullying were 4.72 times more likely to develop anxiety, while those who faced severe bullying were 18.52 times more likely (Li et al.). Physical bullying had the strongest connection to anxiety, followed by verbal and relational bullying. This suggests that as bullying becomes more serious, the risk of anxiety also increases.

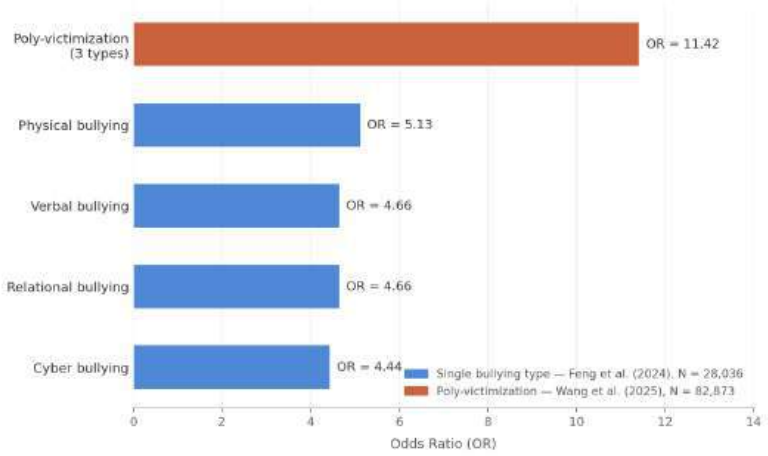


Figure 1. Odds ratios (ORs) of anxiety risk by bullying type among Chinese students (8–19). ORs represent the increased likelihood of anxiety symptoms compared to students with no bullying exposure (reference = 1.00). Single-type bullying data reflect "frequent" exposure. Poly-victimization refers to simultaneous exposure to traditional bullying victimization, cyber-bullying victimization, and cyber-bullying perpetration. All ORs are statistically significant ($p < .05$). Figure generated by author.

Beyond the numbers, many bullied students describe feeling scared and unsafe at school. Some are afraid to tell teachers or parents, which can make their anxiety build up over time. Victims of cyberbullying often feel even more stressed because the bullying can happen at any time, not just at school. One student described the experience by saying, “Almost no one plays with me. I feel I am not as good as the others” (Li and Hesketh). This shows how bullying can damage students’ self-esteem and emotional well-being. This anxiety also affects their behavior.

Some students start avoiding certain places or skipping classes to stay away from the people who bully them. Research also shows that bullying can negatively affect academic performance and make it harder for students to focus in school (Wang et al.). Over time, this can hurt their grades and make it harder to build friendships. It becomes a cycle that is difficult to break.

2. Depression

In addition to anxiety, depression is another major mental health issue linked to bullying in China. The same Sichuan study found that 20.9% of students showed depressive symptoms overall. However, among bullied left-behind children, the rate increased to 57.7%, compared to 22.1% among those who were not bullied (Feng et al.). A nationwide study of 15,415 middle and high school students also found that bullying was strongly associated with depression and other mental health problems (Luo et al.).

The data shows a similar pattern to anxiety. Students who experienced mild bullying—meaning it happened less often or was less serious—were 3.64 times more likely to develop depression, while those who faced severe bullying—which happened more often or was more serious, such as physical bullying—were 5.34 times more likely (Li et al.). At the same time, students who already have depression are also more likely to become involved in bullying. One study found that students with moderate-to-severe depression were more likely to be bullying perpetrators (Wen et al.). This suggests that bullying and depression can form a vicious cycle, where being bullied can cause depression, and depression can also make students more likely to get involved in bullying.

Real-life cases show how serious bullying-related depression can become. In 2016, a 14-year-old boy in Qinghai Province died after drinking pesticide because he could not handle long-term bullying from classmates. In 2023, a 22-year-old university student in Zhejiang died after jumping from a building following bullying from her roommates, even after asking for help (Lin). These cases show how dangerous the long-term effects of bullying can be if no one steps in.

For many students, depression develops slowly. Victims often describe feeling hopeless and losing interest in things they used to enjoy. Left-behind children may be especially vulnerable because they do not have their parents around for support (Feng et al.). Even after the bullying ends, the emotional effects can last for a long time, as studies have shown that the impact of bullying can continue well beyond adolescence (Bai et al.).

3. PTSD

Repeated bullying can make students feel constantly scared and unsafe, which may lead to psychological trauma over time.School bullying is not just temporary unhappiness; it can leave long-term psychological scars (Luo et al.). A large survey of 82,873 students in Sichuan Province shows that bullying victims face a very high risk of PTSD (Wang et al.). Students who experience physical bullying, cyber-victimization, and cyber-perpetration at the same time are about 15.48 times more likely to develop PTSD than others. This suggests that facing multiple

types of bullying can lead to more serious psychological harm, because students are exposed to repeated stress across different environments, such as school and online spaces. This constant pressure makes it harder for them to feel safe, and the negative emotions can build up over time instead of going away. As a result, their mental health becomes more seriously affected. In addition, bullying can affect students' emotional regulation, making them feel more easily angered, depressed, and vulnerable, and these reactions often become stronger as bullying gets more severe (Education Blue Book).

In real life, this kind of trauma often feels like a lasting psychological shadow (Li and Hesketh). Many students say that even a long time after the experience, their hearts still race when they see similar places at school or hear certain voices. This is a common sign of PTSD (Nielsen et al.). Over time, some students stop feeling safe and begin to distance themselves from others. This constant sense of alertness can be exhausting, almost like they are always waiting for something bad to happen. Because of this, it becomes harder for them to focus on daily life and school.

4. **NSSI**

Bullying can create intense emotional pressure and a sense of helplessness, which may push some students toward harmful coping behaviors. When emotional pain becomes overwhelming, some students may turn to self-harm (Li et al.). Data shows that the rate of non-suicidal self-injury (NSSI) among Chinese adolescents is around 22.37% to 26.1% (Chen et al.; Li et al.). For bullying victims, the risk of self-harm is about 2.1 times higher than for others. In some groups, the situation is even more serious. For example, in rural Jiangxi, about 47.3% of left-behind children reported engaging in self-harm. In a study in Changsha, 4.1% of middle school students admitted to this behavior (Chen et al.).Research also shows that mental health problems, such as depression and anxiety, explain about 12.96% of the connection between bullying and self-harm, meaning that most of the effect is direct rather than going through these psychological factors (Li et al.). This suggests that bullying may lead to self-harm partly because it damages students' mental health, but also because the stress and pain can directly push students toward harmful coping behaviors. This suggests that bullying may lead to self-harm partly because it damages students' mental health.

For these students, physical pain can sometimes temporarily distract from emotional suffering (Li et al.). Some victims use self-harm as an unhealthy way to cope, especially when they feel unsupported (Li and Hesketh). In rural areas, left-behind children often lack parental support, so when they are bullied, they tend to keep their feelings inside. Over time, some may turn to self-harm as a way to release that pressure. In many cases, this can be seen as a form of silent help-seeking. If bullying continues without intervention, this feeling of hopelessness can become even more serious.

5. Long-Term Effects

Bullying can have long-lasting effects that continue to influence a person's life even after they leave school. These effects can impact mental health, career development, and social relationships, and can also affect those who bully others.

Research shows that people who were bullied are about four times more likely to develop major depression ten years later (Bai et al.). This shows that the emotional impact of bullying can last for many years. Bullying can also affect career outcomes. Students who are bullied may struggle in school, which can reduce their future opportunities and lead to lower productivity in adulthood (Wang et al.; Liu et al.). In addition, those who bully others are more likely to face long-term problems. They have a higher risk of violent behavior, substance abuse, and criminal activity later in life (Yang et al.). For example, in 2016, 1,180 people in China were arrested for school bullying-related violence (Han et al.).

Beyond the data, bullying can also affect how people think and behave. Many victims become more withdrawn and may have difficulty trusting others, which can make it harder to build close relationships later in life (Li and Hesketh; Yang et al.). Victims may also continue to struggle with low self-esteem, which means they may feel less confident and unsure about themselves in social situations (Li and Hesketh; Feng et al.). At the same time, those who bully others may continue these harmful behaviors, which can develop over time and create a cycle where these behaviors repeat and affect both themselves and others (Yang et al.; Han et al.).

LIMITATIONS

This study has several limitations. First, since this paper is based on existing research, it cannot fully investigate the distinction between correlation and causation. For example, while bullying is linked to anxiety, depression, PTSD, and self-harm, it is not always clear whether bullying causes these problems or if students with these issues are more likely to be involved in bullying.

Second, much of the data comes from self-reported surveys. Students may not always report their experiences accurately because of fear, memory problems, or social pressure. This may affect the generalizability of results.

In addition, many of the studies focus on specific regions in China, such as Sichuan or rural areas. Because of this, the findings may not fully represent all students across different parts of China or other countries.

Finally, this paper primarily focuses on psychological outcomes and factors; other influences, such as personality, school environment, and peer relationships, may also play an important role but are not fully explored in this study.

Future research can help improve this study in several ways. First, long-term research is needed to better understand how bullying affects students over time, particularly into adulthood. Second, researchers self-reported surveys can be supplemented with more quantitative methods, such as fMRI scans and measuring cortisol or serotonin levels. In addition, future studies could include students from more regions to improve the generalizability of findings and remove the

potential for unintended variables. Finally, factors such as personality and school environment require further research to understand bullying more holistically.

CONCLUSION

In conclusion, school bullying in China is a serious issue that affects many students and is closely linked to anxiety, depression, PTSD, and NSSI. These effects can last through a student's education and even continue into adulthood, impacting vocational performance, interpersonal relationships, and even the risk of criminality, meaning that their socioeconomic and psychological impacts are widesweeping. This study also finds that the likelihood that a student participates in bullying is influenced by factors such as repeated exposure to bullying and lack of support from family or school.

Based on these findings, parents and educators should devote more attention to students' emotional well-being, not just their academic performance, as an indicator of readiness for independence and the pressures of a modern workforce. School administrators and families should also create a supportive environment where students feel safe to report bullying and seek help. Additionally, schools can improve supervision and provide more mental health support. Parents should maintain open communication and pay attention to students who may feel isolated. Early intervention is important to reduce long-term harm.

Works Cited

Bai, Yu-Ping, et al. "Longitudinal Study of Peer Bullying Victimization and Its Psychological Effects on Adolescents." World Journal of Psychiatry, vol. 15, no. 5, 19 May 2025, doi:10.5498/wjp.v15.i5.104145.

Chen, Hui, et al. "Influence of Academic Stress and School Bullying on Self-Harm Behaviors among Chinese Middle School Students: The Mediation Effect of Depression and Anxiety." Frontiers in Public Health, vol. 10, 6 Jan. 2023, doi:10.3389/fpubh.2022.1049051.

Feng, Yuan, et al. "Association between Bullying Victimization and Mental Health Problems among Chinese Left-Behind Children: A Cross-Sectional Study from the Adolescence Mental Health Promotion Cohort." Frontiers in Psychiatry, vol. 15, 7 Nov. 2024, doi:10.3389/fpsyt.2024.1440821.

Huang, Lihong, et al. "School Bullying and Its Association with Psychological Wellbeing: Findings of the Fujian Adolescent Mental Wellness Study (FAMWeS)." PLOS ONE, vol. 20, no. 1, 2025, doi:10.1371/journal.pone.0315993.

Li, Hongjing, et al. "Analysis of the Current Situation and Factors Influencing Bullying in Junior High Schools in Backward Areas of Western China: A Case Study of Qingyang City in Gansu." BMC Public Health, vol. 24, no. 1, 13 May 2024, doi:10.1186/s12889-024-18775-5.

Li, Jiameng, and Therese Hesketh. "Experiences and Perspectives of Traditional Bullying and Cyberbullying among Adolescents in Mainland China—Implications for Policy." Frontiers in Psychology, vol. 12, 6 July 2021, doi:10.3389/fpsyg.2021.672223.

Lin, Shengqiao. "An Exploration of the Status, Cases, and Countermeasures of Adolescent School Bullying." Open Journal of Legal Science, vol. 12, no. 12, 2024, pp. 6947–6953, doi:10.12677/ojls.2024.1212986.

Liu, Wenting, et al. "Uncovering the Prevalence, Risk Factors and Consequences of School Bullying in China's Rural Primary Schools." BMC Public Health, vol. 25, 2025, doi:10.1186/s12889-025-25101-0.

Li, Xuyang, et al. "Research on the Relationships between Psychological Problems and School Bullying and Non-Suicidal Self-Injury among Rural Primary and Middle School Students in Developing Areas of China." International Journal of Environmental Research and Public Health, vol. 17, no. 10, 12 May 2020, doi:10.3390/ijerph17103371.

Luo, Xiaomin, et al. "Relationship between School Bullying and Mental Health Status of Adolescent Students in China: A Nationwide Cross-Sectional Study." Asian Journal of Psychiatry, vol. 70, 2022, p. 103043, doi:10.1016/j.ajp.2022.103043.

Nielsen, Morten B., et al. "Post-Traumatic Stress Disorder as a Consequence of Bullying at Work and at School: A Literature Review and Meta-Analysis." Aggression and Violent Behavior, vol. 21, 2015, pp. 17–24, doi:10.1016/j.avb.2015.01.001.

21st Century Education Research Institute. Annual Report on China's Education (2017) [Education Blue Book]. Social Sciences Academic Press , 2017.

Wang, Yu, et al. “The Prevalence and Association of Traditional Bullying and Cyber Bullying with Mental Health among Adolescent and Youth Students in China: A Study after the Lifting of COVID-19 Restrictions.” BMC Public Health, vol. 25, no. 1, 14 Feb. 2025, doi:10.1186/s12889-024-20940-9.

Wen, Yang, et al. “Mental Health, Bullying, and Victimization among Chinese Adolescents.” Children, vol. 9, no. 2, 11 Feb. 2022, p. 240, doi:10.3390/children9020240.

Xing, Jianli, et al. “The Prevalence of Bullying Victimization and Perpetration among the School-Aged Population in Chinese Communities: A Systematic Review and Meta-Analysis.” Trauma, Violence, & Abuse, vol. 24, no. 5, 2023, pp. 3445–3460, doi:10.1177/15248380221129595.

Assessing the Influence of ESG Considerations on Early-Stage Venture Capital Decision-Making By Rishav Mehta

Abstract

The purpose of this study is to analyze the effect of environmental, social, and governance considerations in venture capital decisions. While the importance of ESG practices has been established in institutional investment and public market investments, the application of ESG criteria to venture capital is less evident because of the difficulty in collecting enough data from emerging firms. The present research applied a qualitative exploratory design and used secondary sources of information and semi-structured interviews. Three interviewees were asked to share their views and insights on ESG in venture capital: one of the participants studied at the University of Texas at Austin and had experience in venture capital and private equity, another interviewee is a professional from the field, and the last one is an angel investor and startup advisor. It appears that when it comes to ESG in venture capital, investors perceive it as a qualitative criterion that helps to evaluate risks, the resilience of the company, the judgment of the founders, and the potential for value creation. Governance was identified as the most important ESG dimension by all participants, particularly as related to ethical decision-making, leadership, cap table management, data privacy, and accountability. While stakeholders typically perceived ESG to be applicable in due diligence and portfolio management, there were instances where ESG issues arose in screening, investment committee decisions, and exits. This research finds that, at present, ESG serves as both a risk management and signaling framework in venture capital financing. According to stakeholder theory, signaling theory, and institutional theory, it is believed that the integration of ESG in venture capital is ongoing and may become more structured in the future.

1. Introduction

The last two decades have seen the emergence of Environmental, Social, and Governance (ESG) considerations as a new paradigm that has changed the face of international finance. What used to be perceived as a specialized issue pertinent to socially responsible investment funds is now a key element of the investment strategy of large asset managers, institutions, and private equity companies. The development of a wide array of ESG standards, including the United Nations Principles for Responsible Investment (UN PRI), Global Reporting Initiative (GRI), and Sustainability Accounting Standards Board (SASB), indicates a growing trend of alignment between financial success and sustainable development. As reported by the Global Sustainable Investment Alliance (2022), the share of assets managed under ESG-related strategies increased to over $35 trillion in 2021, representing more than one-third of all globally managed assets.

While many topics within the ESG realm have been covered extensively, there is still one area that is relatively underexplored: the importance of ESG metrics in making venture capital (VC) investment decisions during the early stages of development. Unlike publicly traded securities and late-stage private equity, early-stage ventures are characterized by high levels of

uncertainty and a lack of any historical data. In this regard, traditional VC due diligence is largely based on qualitative criteria, such as the capabilities of the founding team, the size of the target market, and the scalability of the product/service. The application of ESG metrics to early-stage venture investments is rather inconsistent and is often regarded as less important than the opportunities for growth or exit. Nevertheless, the mounting pressure from LPs to adopt ESG principles and the increased understanding that sustainability can be an asset have led VCs to consider ESG at the early stage of their operations.

This study will examine the role played by ESG measures in venture capital investment decisions at an early stage and pay particular attention to the manner in which VC firms around the world incorporate or fail to incorporate ESG into their due diligence process, active monitoring of their portfolios, and value creation. While many streams of literature have addressed the issue of ESG integration in the public markets, there is very little empirical data available about the use of ESG in early-stage investments and high-growth. On the one hand, some researchers claim that ESG helps mitigate operational and reputational risk even among startups (Gompers et al.; Barber et al.). On the other hand, other researchers state that the absence of a common framework for reporting and the pressure on rapid growth make ESG assessment impossible for seed and Series A investments (Cumming and Johan). The core rationale behind the discrepancy between these opinions is to prove the need for a comprehensive analysis of whether ESG measures affect venture capital investment decisions.

The Problem and Research Gap

Whereas ESG integration has become a prevalent practice in institutional investment, research into the effects of ESG on venture capital decision-making is rather limited. The literature has mostly explored the relationship between ESG and financial performance in listed firms, mainly concentrated on large-cap companies with sufficient information available (Friede et al.; Krüger). Such research has found that high ESG performance can lower the cost of capital, draw long-term investors, and improve corporate reputation. Nevertheless, the information setting in venture capital is substantially different from that in public companies. Emerging ventures do not report on sustainability issues and have undeveloped corporate governance structures. Besides, the incentive structure in venture capital, determined by the cycle of fund management and the exit process, might be inconsistent with the long-term nature of ESG considerations.

Thus, venture capitalists are caught in a difficult situation because their institutional investors expect ESG compliance, but the ventures they invest in do not have the capability or the will to provide valid ESG data. This creates an interesting discussion on the way ESG can be applied or disregarded when it comes to early-stage investments. According to a recent PwC report, even though 65 percent of venture capital firms use ESG in their due diligence process, only 20 percent have developed ESG scoring systems.

Purpose and Significance of the Study

The purpose of this research was to critically analyze the process of measuring, prioritizing, and implementing ESG indicators as part of venture capital investment decision-making during the early stages of investment. The research sought to explore whether the ESG indicators were used to influence the investment decision-making process or whether they served as signals to the investors. This was done through reviewing existing literature and interviewing participants in venture capital, angel investing, and start-up companies.

The importance of this topic lies in the fact that while ESG factors have been explored thoroughly in public markets and institutional investing, they remain relatively unexplored in entrepreneurship. Startups, especially in their initial phases, tend not to have any ESG reporting done or a structure in place that can measure their sustainability; thus, ESG becomes somewhat more difficult to measure compared to public companies. The current paper uses input from professionals in the field of venture capital, an angel investor, and a student who is acquainted with venture capital and private equity.

Research Question and Hypothesis

The central research question guiding this paper is:

> To what extent do ESG metrics influence early-stage investment decisions in venture capital, and through what mechanisms are these metrics integrated into the due diligence process?

The working hypothesis suggests that the role of ESG measures in venture capital decision-making is mainly qualitative, driven by considerations such as risk assessment, founder assessment, and LP signaling. Although startup investors might not quantify their ventures' ESG measures, they do consider sustainability factors, including issues of corporate governance, diversity of employees, and environmental impact, as implicit indicators of sustainability.

Conceptual and Theoretical Foundations

The current research is based on the stakeholder theory and the signaling theory. Freeman’s stakeholder theory implies that companies create value when they consider the interests of various stakeholders such as employees, consumers, suppliers, and the community as a whole. In the venture setting, startups that practice stakeholder-based governance can attract investors who believe that they are flexible and ethical. Spence’s signaling theory supports this approach through its implication that environmental, social, and governance initiatives act as signals of managerial excellence and strategic vision, particularly in the situation when information asymmetry exists, which is typical of venture investments.

2. Theoretical Framework and Literature Review

2.1 The Evolution and Definition of ESG Metrics

The development of the Environmental, Social, and Governance (ESG) approach is linked to the larger trend of socially responsible investment (SRI), which became prevalent

during the 1960s and 1970s. The initial socially responsible investment approaches were based on exclusionary screening, meaning that individuals refrained from investing in harmful sectors such as tobacco, military, and environmentally destructive industries. However, it was not until the early 2000s that ESG was formally introduced through the 2004 United Nations report titled "Who Cares Wins," where the business case for integrating environmental, social, and governance considerations was outlined. This led to the development of the United Nations Principles for Responsible Investment (UN PRI) in 2006.

Today, ESG criteria are widely employed for measuring a firm's non-financial performance in three broad categories. The first category is the environmental category, which involves resource efficiency, carbon emissions, waste management, and vulnerability to climate change. The second category is the social category, which includes labor practices, human rights, diversity, and community engagement. The third category is the governance category, which covers board independence, executive pay, transparency, and corruption. As noted by Khan et al., companies with high performance on material ESG criteria, criteria that are significant in their respective industries, generate substantially higher risk-adjusted returns compared to similar companies with poor ESG performance.

However, the actual practice of measuring ESG continues to be inconsistent. Efforts to standardize reporting have been made through frameworks such as the Global Reporting Initiative (GRI), the Sustainability Accounting Standards Board (SASB), and the Task Force on Climate-related Financial Disclosures (TCFD). Yet, discrepancies between the scope and methods used continue to exist. Private and early-stage companies find it even harder to report because they do not have to disclose information on their ESG performance.

2.2 ESG and Financial Performance: Evidence from Public Markets

Much research has been carried out on the correlation between ESG performance and financial results, mainly focusing on public equity markets. A meta-analysis by Friede et al., for instance, which reviewed over 2,000 studies, discovered that almost 90 percent of these indicated a non-negative relationship between ESG and corporate financial performance, with the majority indicating a positive link. In a similar vein, Clark et al. showed that firms with strong ESG performance enjoy advantages such as reduced cost of capital, superior operational performance, and better stock performance over time.

Nevertheless, the processes by which ESG affects financial performance differ from one case to another. For example, the risk reduction theory is based on the idea that organizations that follow stringent ESG guidelines face fewer risks, such as regulatory punishment, reputational risks, and environmental risks (Bassen and Kovács). On the other hand, the resource efficiency theory holds that firms that operate sustainably are able to generate efficiencies and innovative solutions, resulting in higher performance (Porter and van der Linde). Another theory is the stakeholder management theory, which suggests that good governance helps build better relationships with stakeholders, thus leading to profits (Freeman).

Though such empirical consensus offers a robust basis for the importance of ESG in capital markets, it is largely based on big, publicly held companies that can generate sufficient reports and be subjected to investor scrutiny. This distinction becomes very important when one tries to apply ESG in venture capital since startups do not possess adequate information or long-term governance frameworks.

2.3 Venture Capital Investment and Decision-Making Processes

Venture capitalists (VCs) have many differences from traditional asset managers with respect to the investment time frame, decision-making criteria, and willingness to accept risks. As described by Gompers et al., venture capital firms make investments in early-stage firms that involve high levels of uncertainty and information asymmetry, with the decision-making process involving more qualitative assessments than quantitative analyses. The three key considerations in making decisions include the capabilities of the founders, the market potential, and the scalability of the business model.

MacMillan et al. noted that the competence of the founding team was the most important factor that had an effect on the investments made. Further studies have confirmed that soft variables, like teamwork, adaptability, and ethical behavior, have a great influence on investors' confidence (Franke et al.). In light of the above information, the integration of the ESG approach into venture capital does not necessarily require the creation of new models but can fit into the existing ones.

The venture capital process typically consists of three phases: screening, due diligence, and monitoring after investment (Tyebjee & Bruno). ESG integration is possible at any of these stages. For instance, during the stage of screening, ESG can act as a criterion for excluding startups that run businesses incompatible with sustainability standards. At the due diligence stage, ESG can help evaluate risks of operation, the ethicality of the supply chain, or the company's readiness for compliance. Finally, during the stage of monitoring after investments, venture capitalists can rely on ESG approaches to help their portfolio companies scale responsibly. However, there is empirical evidence that ESG integration at these stages is sporadic.

2.4 ESG Integration in Venture Capital

Some studies have recently appeared that provide some insight into how ESG principles are applied within venture capital practice. According to Cumming and Johan, there is a high dependence between the probability of applying the ESG principle and the firm size, as well as the structure of the investment group of the company. It turns out that enterprises that have institutional investors and government organizations apply ESG principles much more often than those financed by individual and family investors.

As highlighted by a study published by the Global Impact Investing Network, almost 70% of venture capital funds are now claiming to be using ESG criteria while evaluating investment opportunities; however, less than one-third actually use a set of standardized

evaluation tools. The reason for the lack of implementation in terms of ESG policies is the challenge of applying consistent ESG metrics to early-stage companies that frequently alter their business models or move into different sectors.

There is also empirical evidence to show that ESG issues related to venture capital have been more concerned with governance rather than environmental and social aspects. According to Krueger et al., investors generally perceive good governance as an important aspect of managing risks and accountability, which is critical for start-up firms. Environmental and social aspects are hard to measure initially and can become pertinent only after the firm starts growing.

2.5 Theoretical Frameworks Informing ESG in Venture Capital

Stakeholder Theory

One of the significant theoretical constructs underpinning ESG integration is stakeholder theory, formulated by R. Edward Freeman. According to this theory, the creation of sustainable value requires a company's capability to balance various stakeholder interests instead of concentrating only on generating profits for the shareholders. Applied to venture management, ventures that show concern for their employees, customers, and corporate governance can expect investor support and ensure sustained growth. Stakeholder theory can also help ventures cope with reputation-related risks and regulatory challenges that are becoming ever more relevant in technology-based markets.

Signaling Theory

The signaling theory, put forward by Michael Spence, demonstrates how companies communicate credible information to minimize the problems caused by asymmetric information in uncertain markets. For new ventures, embracing the ESG approach could be regarded as credible communication on the part of the company about its managerial skills and foresight. The investment community could see the firm's commitment to sustainability as a sign that the management is disciplined, foresighted, and aligned with the latest regulatory and consumer trends. Research by Flamer, for instance, indicates that companies that engage in sustainability projects receive favorable responses from the market.

Institutional Theory

The additional institutional approach takes into account the impact that norms, rules, and expectations may have on organizational actions. Venture capital funds function in institutional settings in which the pressure of limited partners, government regulators, and public opinion affects their decisions. As stated by Suchman, organizations strive for legitimacy through conformity with socially acceptable behaviors. Hence, implementing ESG criteria by venture capital companies can be viewed as a way to maintain legitimacy when dealing with sustainability-oriented stakeholders.

2.6 Empirical Gaps and Implications

While there are theoretical justifications for the use of ESG principles in venture capital investments, empirical evidence is still lacking. Many of the available studies are more descriptive in nature, examining survey responses and self-reporting of ESG strategies rather than quantifiable investment outcomes. There are no long-term studies that examine how ESG impacts the success rate of startups or exit valuations. In addition, few researchers consider the heterogeneity of ESG practices across different industries and startup growth stages.

The lack of data makes room for mixed-method research, where the literature review can be complemented by qualitative interviews. Interviews with investors, analysts, and entrepreneurs will help understand how ESG is perceived and practiced in early-stage investments. Interviews will also help understand if the concept of ESG is being used merely as a symbol to satisfy LPs.

In conclusion, although there is ample theoretical backing for the importance of ESG in relation to financial performance in the literature, little empirical research has been conducted regarding the impact of ESG on venture capital. This research seeks to fill this gap using a qualitative methodology.

3. Methodology

3.1 Research Design

The study used a qualitative, exploratory research approach to examine the impact of ESG indicators on venture capitalists' investment decisions. A qualitative research approach was suitable because it helped gain insights into the behavior and perceptions of investors that could not be fully obtained through quantification. The study aimed to uncover how venture capitalists incorporated ESG indicators into their decision-making processes during the screening, due diligence, and post-investment phases.

The research utilized two different components. First, there was a secondary research component based on analyzing relevant academic literature, industry publications, and case studies on the use of ESG principles in venture capital. The second component involved primary research in the form of semi-structured interviews with individuals from the investment world, as well as entrepreneurship students at the University of Texas at Austin who had been involved with startup funding or sustainability ventures.

The main purpose was not to quantify the financial implications of ESG, but rather to understand how ESG issues were viewed and addressed in decision-making processes. Since early-stage ESG assessment was largely dependent on context, qualitative evidence offered a stronger basis for identifying causal mechanisms and informing future quantitative research.

3.2 Research Questions

The study was guided by the following research questions:

1. How do venture capitalists and early-stage investors define and interpret ESG metrics within their investment framework?

2. At what stages of the venture investment process are ESG factors considered most relevant?
3. What challenges and barriers prevent the consistent integration of ESG metrics in early-stage investments?
4. How do investors reconcile institutional ESG mandates with the practical limitations of early-stage startups?
5. To what extent do ESG considerations influence risk perception and portfolio management strategies?

These questions were designed to uncover both explicit and implicit mechanisms that shaped ESG integration, allowing for triangulation between literature-based insights and primary qualitative findings.

3.3 Data Sources

3.3.1 Secondary Data

The secondary research component consisted of peer-reviewed academic research, consultancy reports, and financial analysis on ESG and venture capital. The main sources of information included the Journal of Sustainable Finance & Investment, Harvard Business Review, consultancy reports from PwC, Deloitte, and GIIN, as well as databases such as PitchBook and CB Insights.

The selection criteria for secondary data involved relevance to ESG incorporation in private equity and venture capital, methodological soundness, and relatively recent publication dates. The secondary sources were assessed based on their credibility, level of analysis, and consistency with the research objectives. They provided a theoretical foundation and helped establish an understanding of existing theories behind ESG adoption by investors.

3.3.2 Primary Data

The primary data was gathered through three semi-structured interviews conducted with participants who had either firsthand or adjacent experience with venture investments, venture capital, angel investing, venture advising, and entrepreneurial finance. This was done intentionally by keeping the set of participants very narrow since the research is specifically concerned with a niche within finance, namely, ESG considerations in venture decisions.
All three individuals portrayed various roles in the startup investment community. The first individual was an investment analyst who had expertise in private equity, venture capital, analysis of startups, and student investment organizations. The second individual was an expert in venture capital who directly evaluated startups. The third individual was an angel investor and mentor of startups with knowledge of advising and analyzing startups.

Semi-structured interviews were conducted as they gave room for consistency and flexibility in the interview process. The participants were subjected to interview questions related

to the research questions of the study, such as ESG definitions, when ESG factors play a role in the investment process, barriers to ESG integration, balancing institutional and startup expectations, and risk perception and portfolio management influenced by ESG.

3.4 Sampling and Participant Selection

The research adopted purposeful sampling, whereby the sample was drawn from among individuals relevant to the research topic as opposed to being drawn using random sampling techniques. Due to the specialized nature of ESG integration in early-stage venture capital, the sample comprised participants with specialized knowledge and experience in evaluating startups and investing in early-stage ventures.

While the initial research methodology expected that more interviews would be conducted, the final sample involved only three participants who provided different perspectives about early-stage investment. The three participants included an investment analyst with venture capital and private equity experience, a venture capitalist, and an angel investor/startup advisor. The small sample size was suitable for the research design because it enabled a meaningful comparison between only three relevant stakeholders' views as opposed to conducting many shallow interviews.

The viewpoint of the investment analyst assisted in providing a context of ESG as perceived by young finance professionals with venture capital and private equity experience. The venture capitalist viewpoint helped understand how ESG is incorporated in formal investment analysis and valuation. The angel investor and entrepreneur viewpoint assisted in understanding how ESG is considered in the initial stage of financing start-ups, which do not have any formal reporting systems. Thus, this combination of viewpoints assisted in exploring the concept of ESG in various stages of financing start-ups.

3.5 Primary Data Collection: Interview Component

Apart from secondary research, three semi-structured interviews were also undertaken in order to obtain primary data about the perception of and approach to ESG in early-stage venture capital. The first interview took place on November 11, 2025, through Zoom with verbal permission to record the conversation. This particular individual had knowledge in areas such as private equity, venture capital, startup analysis, and investments.

Further two interviews were undertaken to expand the scope of the research beyond just one viewpoint of an early-career investor. An interview was conducted with a venture capitalist, and another interview was conducted with an angel investor and startup consultant.
The discussion lasted approximately **40 minutes** and covered six thematic areas:

1. Background and exposure to venture capital,
2. Understanding and definition of ESG,
3. ESG practices in VC and startup evaluation,
4. Attitudes toward ESG's effectiveness,

5. Institutional and market influence, and
6. Future expectations for ESG in venture capital.

3.6 Data Analysis

The interview responses were subjected to thematic analysis, a qualitative research tool employed to examine themes or patterns in the data collected through the process of interviews (Braun & Clarke). Thematic analysis entailed the identification of themes running through the three interviews and not just one-off opinions.
The process involved three stages:

1. **Initial coding:** Identifying significant statements or patterns in participant responses.
2. **Theme development:** Grouping related codes into overarching categories such as "ESG as risk mitigation" or "LP-driven ESG compliance."
3. **Interpretation:** Relating these themes to theoretical frameworks (stakeholder, signaling, and institutional theory) and to findings from secondary literature.

By using both the data obtained from the interviews and secondary sources, the analysis reviewed the way ESG impacts the perception of investors instead of trying to quantify the impact of ESG on the financial results of startups. The choice was justified by the fact that early-stage ESG assessment may be quite subjective and context-dependent.

3.7 Validity, Reliability, and Limitations

Validity was improved due to triangulation of findings from the interviews and literature review. At the same time, the three interviews offered different perspectives regarding early-stage investment environments, such as those of investment analysts, venture capital professionals, and angel investor/entrepreneurial advisor perspectives. Reliability was ensured through semi-structured interview guides that required all three interviewees to address similar issues, but still offered some flexibility in their answers.

One of the limitations of the study was its small sample size, namely, just three interviews. As a result, the findings of the research cannot be applied to venture capital companies, angel investors, particular industries, or any market in general. However, the aim of the study is not to generalize. It is an exploratory study dedicated to a narrow topic.

Another limitation of the research is associated with the fact that it used self-reported data. Being a socially desirable topic, ESG could have been overrated by participants. In addition, the study investigated investor perception rather than investment outcomes and thus cannot prove that ESG integration results in better financial performance of startups.

4. Results

4.1 Overview

The results of this study consist of both secondary research and three semi-structured interviews to analyze how ESG factors manifest themselves in venture capital decisions made at an early stage. The participants of interviews did not consider ESG as a strict score system or a principal replacement of regular criteria used for investments.

Five major findings can be highlighted. First, ESG was perceived differently in early-stage venture capital compared to public markets, which is associated with the lack of formal disclosure requirements, performance history, and reporting standards among startups. Second, ESG became significant for due diligence and portfolio management after investments, but some issues regarding ESG can emerge at the stages of screening and exit preparation. Third, the main obstacles to incorporating ESG include the lack of data and standardization, limited resources, and inconsistency between startup-specific needs and mature company requirements. Fourth, investors justified their ESG expectations for startups based on the stage-based approach, with growing expectations as the company grows older. Fifth, ESG impacted the perception of risks through the identification of governance, reputational, regulatory, privacy, and other issues.

4.2 How Investors Define and Interpret ESG

Based on the interviews, early-stage investors perceived ESG through practical indicators of whether a company could develop sustainably. In the case of public organizations, the assessment of ESG performance typically involves the use of disclosures, third-party ratings, and frameworks. In the case of startups, participants noted that ESG is mostly assessed based on the behavior of founders, culture, riskiness of the business model, and trajectory of the company. A representative of the venture capital sector defined ESG as metrics regarding risk, resilience, and long-term value creation. The participant stressed that early-stage investors do not expect startups to provide comprehensive reports regarding ESG performance. They would rather focus on the fact that founders comprehend the long-term implications of their actions. Relevant aspects include resource efficiency, supply chain exposure, data privacy, employee treatment, structure of the cap table, IP ownership, and decision-making process.

In contrast to the venture capital participant, the angel investor viewed ESG more pragmatically as a way to protect investor capital and identify companies ready for growth. In this regard, the framework can be useful since it allows detecting if a company has proper leadership, is ethical, and has sufficient organizational discipline to escape possible future problems.

Finally, the investment analyst participant pointed out that, in the venture capital industry, the approach to ESG is informal and is strongly related to governance. As the participant stated, ESG can be viewed as both substantive and symbolic in terms of its nature.

4.3 Stages Where ESG Matters Most

There seemed to be a consensus that ESG was the most important for due diligence and support after investments were made. In the context of due diligence, ESG was used by investors to spot potential risks associated with a startup’s scalability. These risks could include vague

ownership structure, lack of good governance, data privacy concerns, questionable marketing practices, bad HR practices, regulatory risks, and potential reputational risks.

In the case of a venture capitalist participant, ESG showed itself to be present in multiple stages of the investment process. At the stage of sourcing and screening, ESG allowed avoiding sectors or types of businesses that had clear environmental, social, or regulatory risks. At the stage of due diligence, ESG got more attention as it consisted of the evaluation of the founders' history, ethics of the products, potential risks associated with the supply chain, potential problems with the use of AI and data, and governance issues. At the stage of the investment committee/term sheet, ESG could show up in light form as expectations for reporting progress, ethics, or diversity.

In addition to discussing the investment period, the angel investor discussed the phases of pre- and post-investment as well. In the pre-investment phase, ESG was helpful in weeding out companies with big red flags. In the post-investment phase, ESG played a role in guiding founders on matters of governance, reporting, cybersecurity, and sustainability. From these results, it can be seen that ESG does not only have an impact during one phase of the investment process.

4.4 Barriers to Consistent ESG Integration

All three individuals cited significant barriers that impede the consistent application of ESG to early-stage investing. The most popular barriers included limited availability of data, lack of standardization in measurements, resource limitations, and complexity of using the ESG framework designed for mature companies to assess startups.

The venture capitalist said that startups generally do not have the infrastructure to measure their ESG performance. Most of them are still preoccupied with establishing product-market fit, acquiring customers, recruiting employees, raising funds, and even staying alive. Therefore, extensive ESG reporting is unrealistic at an early stage. Also, ESG materiality is a challenge because startups pivot frequently, enter new markets, or change their products. The angel investor agreed with the above viewpoint, saying that startups often lack the resources to measure ESG formally. ESG might be relevant, but it cannot be applied to a seed-stage company in the same way as it can be applied to a mature publicly traded company. The investment analyst also pointed out the challenges related to the measurement of ESG, as it is complicated to measure ESG in startups because of limited data availability.

These results imply that the issue is not with the reluctance of investors to engage with ESG. It is more about the implementation of such an approach, despite the recognition by investors of its applicability to small, volatile, and dynamic firms.

4.5 Stage-Based ESG Expectations

It became evident from the interviews that many investors reconcile ESG expectations and startup realities with a proportional stage-based approach. Namely, while not expecting

early-stage startups to have the same level of ESG compliance as mature companies, they adapt expectations based on the company's size, investment stage, revenue, team size, and industry. This idea was referred to by the venture capitalist interviewee as a “principle over process” approach. Instead of requiring young companies to go through an audit and ensure compliance with all requirements, they focus on the founder’s intentions and ethics, responsible product design, and consideration of risks associated with business operations. Further ESG expectations can then be established when the company matures.

The angel investor expressed a similar opinion about the importance of a “grow with the company” approach. During the early stages, investors may only expect some basic practices such as proper corporate governance, diversity awareness, basic cybersecurity, and ethical leadership. As the company raises further rounds and matures, expectations can become more sophisticated.

Thus, the results of this investigation demonstrated that investors do not necessarily see a conflict between ESG and company growth.

4.6 ESG, Risk Perception, and Portfolio Management

Risk perception was one of the areas participants constantly associated with ESG. Participants stated that ESG was a means to understand regulatory, reputational, operational, and governance risks that could have a detrimental effect on the startup’s value in the long run. In this regard, the venture capitalist noted that due to a lack of resources and reputation, the impact of the issues related to governance risks, regulation risks, data risks, AI ethics, and public risks could be detrimental to startups since they are fast-growing enterprises without any experience. In turn, the angel investor linked ESG to the prevention of liability, suggesting that good leaders and their commitment to ethics, cybersecurity, diversity, and sustainability could help to identify whether the startup is ready to scale.

Furthermore, ESG could be used in the process of portfolio management by investors, who could take it into account when supporting the companies in their portfolio, making follow-on funding decisions, preparing the startups to further fundraising and acquisitions. A good ESG story could help to make a company more attractive to investors and partners.

4.7 Cross-Interview Synthesis

In all three interviews, ESG emerged as an informal, but increasingly relevant factor for early-stage venture capital investments. The venture capitalist interviewee offered the most structured point of view, viewing ESG as a concept that may be present throughout screening, due diligence, investment committee discussion, post-investment management, and exit strategy. The angel investor discussed the very early stages of business creation and talked about capital preservation, the maturity of founders, and their gradual ESG development. The investment analyst gave an insight into the practical and symbolic aspects of ESG, viewing it as not only an actual ethical concern but also a credibility indicator.

Overall, these results confirm the working hypothesis of this study. Early-stage investments are affected by ESG mostly in terms of qualitative rather than quantitative factors. ESG impacts the way investors evaluate risks, credibility of founders, operational maturity of companies, and their readiness for future operations. Yet, ESG is hardly ever decisive when compared to more classical venture capital factors, such as market potential, scalability, product quality, and strength of the founders' team.

5. Discussion and Conclusion

5.1 Interpreting the Findings

The results indicate that ESG in early-stage venture capital should be considered an emerging investment logic rather than a concrete investment framework. The investors see the relevance of ESG, but very rarely do they score it formally at the very first stage of the company's formation. In the context of early-stage investing, ESG becomes a framework for risk assessment, an entrepreneur's judgment, a company's maturity, and its prospects for growth. Such a conclusion can help clarify why ESG in venture capital differs from ESG in public markets. The companies listed above have the data about their activities, are rated, have a certain history, and have a system of corporate governance. Start-ups, however, lack all these characteristics; thus, their ESG analysis depends more on subjective judgments. Such judgments include entrepreneurs' ethics, the responsible development of the company, neglecting critical risks, and the creation of liabilities.

The study also reveals that the relevance of ESG differs according to the type of investors. The venture capitalist participant perceived ESG in the context of a formal process of investment, the angel investor perceived ESG through the lens of capital protection, while the investment analyst perceived ESG in terms of both ethics and signaling. This implies that ESG is not a homogeneous concept in venture capitalism. Different perceptions of ESG depend on the investor's position.

5.2 Stakeholder Theory and the Role of Responsibility

Stakeholder theory further explains why the relevance of ESG persists despite an absence of an ESG framework among startups. According to stakeholder theory, organizations create more value in the long run if they consider the interests of their stakeholders, like employees, customers, investors, suppliers, and local communities, among others (Freeman). At the early stages of venture capital, however, this is not expressed in any form of comprehensive ESG reporting but rather through actions of the founder and organization development decisions. Such interpretation is supported by the findings of the interviews conducted. The importance of ethical leadership, employee treatment, diversity, reputation among customers, data privacy, cybersecurity, and community engagement was considered to indicate a company's capability to scale sustainably. The reason behind that was that startups that ignore the needs of their

stakeholders may end up facing certain reputational risks, legal action, problems with retention of employees, or loss of customers.

The approach suggested by the interviewees, which involved the development of certain practices depending on the organization's growth phase, also resonated with stakeholder theory. The investors were not expecting early-stage companies to have comprehensive ESG frameworks but rather be aware of their stakeholder-related responsibilities and develop better practices with their growth.

5.3 Signaling Theory and the Value of Reputation

The signaling theory, therefore, becomes relevant to ESG and venture capital in the early stages of ventures because investments in start-ups carry uncertainties and asymmetry of information. In many cases, investors lack information on financial history, proven business models, and official disclosure when evaluating start-ups. Therefore, signals become a means to assess the skills, professionalism, commitment, and risk awareness of the founders.

ESG can serve as one such signal. ESG factors like good corporate governance, clear ownership structure, good data management practices, and stakeholder awareness may all be used to indicate that the founder thinks through and is ready to grow. In cases where ESG is not the basis of an investment decision, it may affect the decision-making process indirectly by decreasing the level of uncertainty about the founder’s competence.

However, it is also applicable to venture capital firms themselves as well. Firms that claim to have an ESG approach in mind may attract the interest of their limited partners, institutional investors, and founders who want to practice responsible investing. However, according to the interviews, ESG signaling is not complete if the firm uses only ESG language and lacks real systems to help its portfolios.

5.4 Institutional Theory and Legitimacy Pressures

According to institutional theory, the reason for the increasing attention paid to ESG in venture capital lies in the nature of institutions, even though the application of such a theory to startups poses challenges. Financial organizations function in an institutionally framed environment that includes limited partners, regulatory bodies, universities, societal expectations, endowments, pension funds, and culture at large.

This explanation fits the study findings quite well. According to the participants of interviews, the introduction of the ESG framework was partly a result of institutional requirements and expectations of limited partners. In other words, ESG in VC is not only the result of a bottom-up approach based on the concerns of founders and individual investors but also a consequence of top-down pressure from capital providers.

However, the pressure from institutions does not necessarily lead to the successful implementation of ESG practices. According to the findings, the early stage implementation of ESG practices takes place on the borderline between legitimacy and practice. The funds introduce ESG practices in order to comply with institutional requirements, but the portfolio

companies do not always implement the framework due to the absence of data, personnel, and reporting systems.

5.5 Rethinking the Small Sample Size

The sample size was small, but this must be seen against the backdrop of the exploratory nature of this study. It did not seek statistical generalizability of the data collected through the interview. The study aimed to collect focused information from participants occupying different positions within the early-stage investment environment.

The three interviewees offered distinct but complementary views: the first participant was an investment analyst with experience in both private equity and venture capital, the second one was a venture capitalist, while the third one was an angel investor and startup adviser. Such organization of the study helped to understand how the perception of ESG changes depending on the stage and position within the startup investment process.

It is essential to mention that the study was not a failure due to its sample size. This was a qualitative study seeking to identify some preliminary patterns in this narrow subject matter. Therefore, the results must be interpreted in the context of the limited number of interviewees, who represented this specific type of population.

5.6 Implications for Venture Capital

These insights raise several considerations for the practice of venture capital investing. First, ESG criteria applied to young companies must necessarily be stage-based; one cannot expect seed-stage companies to fulfill the same ESG criteria as well-developed public companies. Thus, investors should first have minimum expectations regarding governance, ethics, data privacy, and founder responsibility, while the formal ESG reporting will come afterwards.

Secondly, governance must be the starting point for ESG integration into companies' activities. The fact is that the area of governance is the most observable and tangible of all ESG aspects in early-stage startups since it relates to founder decisions, ownership, legal clarity, communication, hiring, and risk management. As for environmental and social aspects, they also remain relevant, but they differ depending on the industry and materiality.

Thirdly, investors should approach ESG through the prism of value creation rather than only compliance. In such a way, ESG may help startups to minimize risks, attract talent, earn customers' trust, get ready for future fundraising rounds, and achieve successful exits.

5.7 Limitations

However, this study still has limitations regarding its small sample and the use of a qualitative design. It is impossible to generalize the findings to other venture capital firms, other industries, and angel investors. The selected sample group gave relevant information; however, it is not statistically significant. Moreover, self-reports were used in the study. It may affect the

results due to the presence of social desirability bias related to the positive perception of ESG in professional circles.

Finally, the study considered investor perceptions, not the outcomes. As a result, it is unable to show whether there is a correlation between ESG integration and better performance, higher survival rates, fundraising success, or exit values of the startups. Future studies should consider a larger sample of interviews of venture capital partners, associates, angel investors, founders, and limited partners. Besides, quantitative studies can be conducted to reveal correlations between particular ESG factors and startup performance.

5.8 Conclusion

It was found that the consideration of ESG factors affects the process of venture capital decision-making at the stage of investment through qualitative evaluation rather than scoring. The investors evaluate riskiness, maturity of the founder, governance, stakeholders' consciousness, and readiness for the future on the basis of ESG principles. ESG applies to due diligence and portfolio management after investment, but it may affect the processes of screening, investment committee discussion, follow-on decision-making, and exit planning.
The key limitations to incorporating ESG into early-stage venture investments include insufficient data, non-standardized measurements, limited resources, and the challenge of using tools that apply to mature companies when dealing with startups. The barriers are overcome through the use of proportional expectations for the development of ESG practices.
The theoretical framework provides an opportunity to understand ESG practices in the context of early-stage investment. Stakeholder theory defines ESG practices as the responsible development of the company. Signaling theory helps to explain the role of ESG in terms of conveying credibility and the maturity of the founder in the situation of uncertainty in investment decision-making. Institutional theory explains ESG adoption in the context of the influence of LPs and other financial practices.

To sum up, ESG in venture capital investments at an early stage is not purely symbolic and at the same time is not yet standardized. It lies somewhere in between. This is an evolving tool that allows identifying risks, establishing trust, and facilitating sustainable scaling. With increased institutional pressure and the development of ESG for startups, it will probably become increasingly standardized in practice.

Works Cited

Barber, Brad M., et al. “Impact Investing.” Journal of Financial Economics, vol. 139, no. 1, 2021, pp. 162–185. https://doi.org/10.1016/j.jfineco.2020.07.008.

Bassen, Alexander, and Anett Kovács. “Environmental, Social and Governance Key Performance Indicators from a Capital Market Perspective.” Zeitschrift für Wirtschafts- und Unternehmensethik, Vol. 9, No. 2, 2008, pp. 182–192.

Braun, Virginia, and Victoria Clarke. “Using Thematic Analysis in Psychology.” Qualitative Research in Psychology, vol. 3, no. 2, 2006, pp. 77–101.

Clark, Gordon L., et al. From the Stockholder to the Stakeholder: How Sustainability Can Drive Financial Outperformance. University of Oxford and Arabesque Partners, 2015.

Cumming, Douglas J., and Sofia Johan. “ESG and Institutional Investment.” Oxford Review of Economic Policy, vol. 35, no. 3, 2019, pp. 553–575.

Flammer, Caroline. “Corporate Green Bonds.” Journal of Financial Economics, vol. 142, no. 2, 2021, pp. 499–516.

Franke, Nikolaus, et al. “What You Are Is What You Like: Similarity Biases in Venture Capitalists’ Evaluations of Start-Up Teams.” Journal of Business Venturing, vol. 21, no. 6, 2006, pp. 802–826.

Freeman, R. Edward. Strategic Management: A Stakeholder Approach. Pitman Publishing, 1984.

Friede, Gunnar, et al. “ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies.” Journal of Sustainable Finance & Investment, vol. 5, no. 4, 2015, pp. 210–233.

Global Impact Investing Network. Annual Impact Investor Survey 2021. GIIN, 2021.

Global Sustainable Investment Alliance. Global Sustainable Investment Review 2022. GSIA, 2022.

Gompers, Paul, et al. “How Do Venture Capitalists Make Decisions?” Journal of Financial Economics, vol. 135, no. 1, 2020, pp. 169–190.

Khan, Mozaffar N., et al. “Corporate Sustainability: First Evidence on Materiality.” The Accounting Review, vol. 91, no. 6, 2016, pp. 1697–1724.

Krueger, Philipp, et al. “The Importance of Climate Risks for Institutional Investors.” The Review of Financial Studies, vol. 33, no. 3, 2020, pp. 1067–1111.

Krüger, Philipp. “Corporate Goodness and Shareholder Wealth.” Journal of Financial Economics, vol. 115, no. 2, 2015, pp. 304–329.

MacMillan, Ian C., et al. “Criteria Used by Venture Capitalists to Evaluate New Venture Proposals.” Journal of Business Venturing, vol. 1, no. 1, 1985, pp. 119–128.

Porter, Michael E., and Claas van der Linde. “Toward a New Conception of the Environment-Competitiveness Relationship.” Journal of Economic Perspectives, vol. 9, no. 4, 1995, pp. 97–118.

PwC. ESG in Venture Capital: Bridging Policy and Practice. PwC Global, 2023.

Spence, Michael. “Job Market Signaling.” The Quarterly Journal of Economics, vol. 87, no. 3, 1973, pp. 355–374. https://doi.org/10.2307/1882010.

Suchman, Mark C. “Managing Legitimacy: Strategic and Institutional Approaches.” Academy of Management Review, vol. 20, no. 3, 1995, pp. 571–610.
Tyebjee, Tyzoon T., and Albert V. Bruno. “A Model of Venture Capitalist Investment Activity.” Management Science, vol. 30, no. 9, 1984, pp. 1051–1066.
United Nations. Who Cares Wins: Connecting Financial Markets to a Changing World. UN Global Compact, 2004.

Why the Separation of Church and State is Necessary for Democracy to Continue
By Violet Day

As of December 25th, 2025, the government passed the official "Religious Division Act," which states that any reference to religion in a state- or federally-sanctioned place is grounds for arrest. To execute the government's will, the Bureau of Enforcement and Secular Tracking (BEAST) was formed. Places where religion is banned are the following: sidewalks, roads, crosswalks, fields, parks, restaurants, schools, the sky, and more.

Almost overnight, the entire nation transformed. Religious buildings were no longer exempt from property taxes, and thus many quickly fell into economic strife. BEAST would take these buildings as their home bases, holding meetings inside the unused chapels, as it would be such a shame to waste them. They declared that everyone had to swear an oath on a stack of state-issued "Secular Code of Conduct" manuals, which would become required reading in curricula nationwide. To further ensure this religious tolerance, BEAST would infiltrate public spaces as a common man to spy on any person who referenced religion. A BEAST missionary once tackled a man for muttering "holy cow," only to apologize afterward because cows, technically, are secular. However, people began to identify the BEAST members because they all had a mark on their right wrist, nicknamed "the mark of the BEAST."

People thought this was the end of their pursuit, but no, it was just getting started. The government forcibly confiscated everyone's money, as the phrase "In God We Trust" was no longer acceptable. When the cash was returned, people found a blank spot where it used to say "God," so now it was "In We Trust." Street names were changed because things like "St. Mary" could no longer be used. Even ones as simple as "Noah" or "Peter" had to be changed. Unfortunately, Google Maps was unable to change the names within the one-hour timeframe, so they were shut down. "Religious promotion is not tolerated in any way, shape, or form," says Delaware Senatorial representative Solomon Faithborne, "even if it's something as simple as a name." What followed was the "Great Renaming," where people had to rename themselves and their children if they had any religious connotations, leading to the near extinction of names like "John," "Caleb," and "Michael." The pushback to these changes was immense, but the people were quickly silenced, leading to a mass influx of underground groups to practice in secret. Despite the negativity, the government considered it a success; the law is functioning as it should: separating church and state in every facet.

Is freedom meant to be seen as a totality, or, as many today would believe, only applied in certain facets? Should we limit freedom when it comes to the laws that govern us and the reasons for their creation? If we must follow the creed of one belief system for the governance of our nation, is that truly freedom in every sense of the word? Many would argue no, but in this day

and age, where reproductive, LGBTQ+, and education topics are being contested, this sense of freedom has become a topic of dispute. In these cases, people often cite religion as their reason for holding a particular opinion. While everyone is entitled to their own view, this becomes an issue when they seek to make their religious opinion law. To bind the law to the belief of one is to go against the very constitution itself; only through a firm separation of church and state can we keep our promise of liberty for all.

At its constitutional foundation, the most prevalent reason is the need to protect individual rights and maintain a fair government. When America was first being established, we were quick to raise this issue and fix it. As the First Amendment says, "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof." This shows that our founders deliberately made sure to cite this necessary distinction from the very beginning to prevent tyranny from taking hold. It's also worth noting the fact that they chose to leave out the word "God" from this description. They did this because, as history shows, whenever you combine religion and government, it ends in disaster. People like the Puritans, who came to the New World to freely practice their faith, were known to persecute non-believers and heretics. Back in Europe, several religious conflicts between the Catholics and Protestants seemed never-ending. King Henry the Eighth himself changed England's whole religion for his own benefit; a direct abuse of power. "The founders had seen the dangers of religious tyranny in Europe," says Dr. Michael Kryzanek, a former professor at Bridgewater University with a PhD in political science, "and sought to prevent similar abuses in the new nation." However, as stated before, certain topics strike a serious chord within religion that cannot go away, which leads to disputes about laws where religious bias comes into play. If said religious bias were to take hold, then our nation would no longer be the land of liberty we claim it to be.

Historically, this is not the first time, nor will it be the last, that this topic has been questioned and discussed. From our nation's founding figures to modern courts, they have contested this same thought, but all have come to the same conclusion. In 1802, Thomas Jefferson spoke of his interpretation of the First Amendment, stating the need for "a wall of separation between church and state." In James Madison's "Memorial and Remonstrance Against Religious Assessments," published in 1785, he argues that government and religion become corrupt when they are intertwined. As seen, these are not new ideas being put forth. They've been present since the foundation of our country. These same ideals are still being carried onward in Supreme Court cases like "Engel v. Vitale" in 1962, which banned school-sponsored prayers in public schools, or "Lemon v. Kurtzman" in 1971, which specified that even indirect government support of religion risks blurring the vital line between faith and law; a line the Constitution is designed to protect. Many of the people pushing these principles were religious themselves, for example, Southern Baptist minister Barry Howard, who states, "The founders understood that mixing the power of government with the authority of religion endangers everyone's freedom." Although he is religious himself, this displays that it is not a matter of restricting faith but expanding freedom for all.

In practice, it's important to note that although these court cases allowed these things to be written into law, they took longer to be implemented. For example, my father, who was born in 1961 and grew up in Alabama, continued the practice of group prayer in their announcements every morning throughout his entire school career until college. However, the people who would lead it would be students or parents, a convenient loophole. In this sense, the government did its job in correcting and clarifying where needed; it's just that the people who oppose finding ways around it will always exist in any form of government that grants freedom. This speaks to the larger, insolvable problem of enforcing such a law once it is put in place across the board.

As many historical figures have pointed out, not only is it unconstitutional, but regular people can get caught in the crossfire, losing their freedom to follow their beliefs, or none at all if they so choose. When you don't correspond with the dominant religion, you will stand out; there's no way around it. You can feel uncomfortable when placed in situations involving religious practices or beliefs that don't align with your own. I myself am agnostic and have felt this sense of alienation from my peers, even without them doing anything to demean me. There have been times, though, when some of them have tried to convert me or assert their moral high ground because of their religious affiliation. While dealing with everyday situations like this can be difficult but doable, it can be even more challenging when the government goes against you as well. Then it turns from a personal disagreement to a matter of the law. To base a law on one religion's moral code restricts another person's rights or choices. For example, what if lawmakers were to prohibit tattoos because some Christian circles interpret the Bible as saying it is wrong? That is unfair to those non-Christians, or even Christians, who believe it is okay to get tattoos. It restricts them from something they have no problem with, therefore restricting their freedom as well. This same principle applies to the contested matters of reproduction, LGBTQ+ rights, and education that we are dealing with today. As journalist David Callaway states, "The government may not favor one faith over another, or religion over nonreligion, because doing so violates the very freedom it is supposed to protect." That's why this separation is vital, not just constitutionally, but for everyone's personal freedom of expression.

Beyond personal experience, others challenge even the legal underpinnings of separation, arguing that the common understanding of the phrase is a misinterpretation. The phrase itself, as legal expert Michael McConnell points out in his interview with Fox News, "is not in the Constitution… I think this is a shorthand version of what the Establishment Clause means." He later says that the true meaning of the phrase is that life and politics should be completely separate, including in public spaces. This is defended by citing Jefferson's original context, noticing that it was raised in the first place to protect religion from government interference. This is still incorrect. Indeed, it was first spoken of to protect religion from the government, but it also works the other way around. This way, freedom is ensured for people on both sides. This, again, is not going to affect your personal relationship with your religion, as that has and will always be protected. It just ensures that others can do the same.

Despite this legal and historical consensus, some critics argue otherwise. As one commentator argues, "Neither the Bible nor our Constitution supports separation of church and

state" ("A Big Misunderstanding"). He states that these laws do more to restrict religious practices, thereby entrenching the spiritual freedom they are said to protect, cementing this idea when he says, "Fear of criticism, reprisal, or persecution is real… Both clergy and everyday Christians come under negative public scrutiny when they speak out." However, this separation marks the difference on the governmental level, not a personal one. When it comes to things like abortion or LGBTQ+ rights, they are not actively forcing those who oppose them to participate in such things; they just allow it to exist lawfully. To outlaw it for religious moral reasons would be unfair to those who do not follow the same religion. So yes, when you speak out on political issues, claiming a moral high ground, you are bound to receive negative feedback, because policy cannot be grounded exclusively in religious doctrine while still serving a diverse public.

Many people following this line of belief become worried that it is encroaching on their sense of freedom. Still, this separation is not to affect your personal relationship with your religion. Several of the founding fathers themselves expressly believed that religion was a matter of privacy, and this separation does not affect your private affairs. When it comes to interacting with those you do not share beliefs with, you handle it situation by situation. As stated before, I am agnostic, and many of my friends are Christian. While some of them have struggled to maintain that respect, as we've gotten older, the majority don't force me to participate in things I do not wish to, like praying before eating. I try my best to never speak down to their religion but rather discuss with them our differing viewpoints. If everyone practiced the same kind of religious respect with one another, then separation of church and state would be seen by all as the simple solution to keeping liberty ever-present.

However, where is this idea of it infringing on religion in public spaces coming from? Surprisingly, it's not entirely unfounded. There have been cases such as the attempt to ban a cross-shaped war memorial, which the court ruled could remain because of its historical significance. This conflicts with the recent law requiring that the Ten Commandments be posted in every classroom in Texas. They used the same historical reasoning here, but there's one key difference: context. For the war memorial, which is a tribute to those who have died, meaning that someone's religious iconography is allowed to be displayed by the families of the deceased. Putting the Ten Commandments in every classroom, on the other hand, serves no greater purpose than just existing there, especially in classes such as math or science. You would not use it for historical meaning in those classes, meaning it does nothing more than promote one's religious beliefs, directly contradicting the separation of church and state and the First Amendment.

At the extremes of this debate, those on the pro-separation side will begin to try to enforce separation into every aspect of our lives, as with the cross-shaped war memorial, which is just not realistic in a world where religion has, and will always, exist. As with any extremist view, this attracts immediate pushback from the other side, as with the Ten Commandments law, and deepens the greater tensions we see today. This is not to fault these outliers, as their intentions are good. Still, these disputes over minor symbols can distract from more pressing civil rights issues, especially concerning abortion and LGBTQ+ topics. Therefore, while they wish to admit it or not, we and the anti-separationists want the same thing: to be able to exist as

we please without infringement on our rights and freedoms, meaning that the separation of church and state is the only logical way to meet both our demands.

But empirically, how widespread is opposition to the separation of church and state? Well, according to the Pew Research Center, "more than half of U.S. adults (55%) express clear support for the principle of separation of church and state when measured this way." In contrast, "roughly one-in-seven U.S. adults (14%) express support for a 'church-state integrationist' perspective." That being said, 18% are mixed, and 12% have no opinion on the topic. As would be imagined, the vast majority of these integrationists are White, southern Christians, showing the commonality between the anti-separationists. This demographic has led to the stance being taken that this law will somehow persecute Christians, as many of this group are quick to harp on the idea, often influenced by apocalyptic or end-times interpretations that heighten their concerns. So if it is such a small populace, why is this argument becoming a highly debatable issue in today's politics? Although this group is numerically small, it remains highly vocal and politically influential, composed of those who usually hold a higher social standing in America's social stratification. This gives them the ability and the means to make their presence known, even if they are not the majority.

Overall, a nation that governs with one hand on the scripture and the other on the law will lose sight of its promised freedom. To maintain our religious equilibrium, we must keep the two separate, as otherwise it will only lead to corruption. As has been shown and defended throughout our nation's history, doing so will keep our government fair and protect those who follow all different kinds of faith, from Christian to Muslim to Atheist. This will further protect those whose rights could be lost if this separation were dissolved. To remain truly free, the United States must continue to draw that sacred line between belief and rule, allowing every citizen to worship–or not–as their conscience guides.

Works Cited

Callaway, David. “What Is Separation of Church and State?” Freedom Forum, www.freedomforum.org/separation-of-church-and-state. Accessed 12 Oct. 2025.

Creitz, Charles, and Michael McConnell. “Constitutional Expert on ‘Separation of Church and State’: Framers Said Nothing Wrong With Religion in Culture.” Fox News, interview by Mark Levin, 13 May 2024, www.foxnews.com/politics/separation-of-church-state-levin-michael-mcconnell. Accessed 9 Nov. 2025.

Howard, Barry. “Five Reasons Separation of Church and State Is Essential – Baptist News Global.” Baptist News Global, 17 July 2023, baptistnews.com/article/five-reasons-separation-of-church-and-state-is-essential. Accessed 12 Oct. 2025.

Kryzanek, Michael. “The Doctrine of Separation of Church and State.” Bridgewater State University, 30 May 2023, www.bridgew.edu/stories/2023/doctrine-separation-church-and-state. Accessed 12 Oct. 2025.

Mitchell, Travis. “In the U.S., Far More Support Than Oppose Separation of Church and State.” Pew Research Center, 12 July 2024, www.pewresearch.org/religion/2021/10/28/in-u-s-far-more-support-than-oppose-separation-of-church-and-state. Accessed 9 Nov. 2025.

“Separation of Church and State: A Big Misunderstanding - the Warrior's Journey®.” The Warrior’s Journey®, 1 Nov. 2021, thewarriorsjourney.org/ethos/separation-of-church-state-a-big-misunderstanding. Accessed 9 Nov. 2025.

Gold Mining and Its Effects on Cocoa Production in Ghana By Leo Cheng

Abstract

Cocoa production is an integral part of Ghana's economy and role in the global system. Tree crops such as cocoa and oil palm are important to smallholders' livelihoods and the national economy (Edberhard 22). However, the cocoa sector is volatile and is massively affected by tradeoffs in land use with gold mining operations (Kolavalli 11). Specifically, it is crucial to determine the extent to which the decline of cocoa exports is based on gold mining in comparison to climate change, land use conflicts, and failures in local governance or the mix of all of them.

Specifically, enforcement has failed in Ghana which has enabled Chinese mining to displace cocoa farming, but other factors like climate change and corruption have made cocoa production decline in the long run as well. Therefore, research needs to be conducted regarding whether or not cocoa mining is the biggest contributor to damage to the cocoa industry, and how much of that problem is directly related to Chinese illegal gold mining, or ASM (artisanal small-scale mining).

Therefore, this study aims to determine how Chinese funded private gold mining operations in Ghana affects cocoa production.

Literature Review

This study aims to address a lack of existing literature by attempting to directly connect cocoa production with foreign mining activity through identifying and qualifying an inverse relationship that is applicable to the entirety of the region in contrast to case studies on smaller ones (Baddianaah 2022).

The broad topic is looking at how the foreign involvement and extraction of gold resulting from increased globalization interacts with the traditional agricultural economies that have shaped Western African countries such as Ghana for most of its history (Kolavalli, Shashi, Marcella 2011). Chinese involvement in Africa but specifically illegal gold mining in Ghana and the Ivory Coast has greatly increased in the last few years which is why current research is either outdated or inapplicable to changing geopolitical tensions and their effect on key industries like cocoa.

Existing scholars focus on China's rising global role in the lens of initiatives such as the Belt and Road Initiative, which is aimed at building peaceful diplomatic relationships with developing countries like Ghana (Park 2023), but lack coverage regarding illegal mining (galamsey) that the Chinese state has turned a blind eye to because of rising demand for foreign extraction (Brooks and Wohlforth, 2016). The most recent literature looks at the effect of Chinese presence on the environment and local governance (Merem 2021), but there is nuance to geopolitical tensions between Chinese nationals and Ghana that is uncontextualized by current literature that helps determine effects on cocoa. Furthermore, there is strong existing research on declining cocoa production due to a litany of warrants like climate change and land usage (Kongor 2024) but it does not delineate how much of changes are caused by which factors,

which contributes to uncertain information about a causal link between Chinese nationals and cocoa production.

Adjacently, a core theme in this research is that China continues to rise as a global power and is building trade and diplomacy with developing nations via symbiotic resource extraction (Park 2023), but motivation for resources like gold has made extractions parasitic in some contexts (Brooks and Wohlforth 2016). Another major theme is that despite cocoa's importance for Ghana and the Ivory Coast's economy, current decline has been met with insufficient current action to reverse the loss of cocoa production.

While sources like (Park 2023) and (Dollar 2016) agree China is helpful for Africa on net, other sources believe they aren't (Merem 2021). Specifically, some scholars would agree that extraction is a net benefit because it reduces poverty and helps develop countries like Ghana with Chinese foreign direct investment that traditionally yields better returns than the US Dollar (Dollar 2016). They also argue that mining can align with cocoa extraction and even boost productivity under certain conditions (Marsh 2015). On the other hand, critics of Chinese involvement cite Chinese state involvement with corrupt local governance systems as the cause of damages (Crawford 2017), while others believe Chinese miners are wholly responsible for damages as an intrinsic byproduct of being involved in the region (Asumang-Yeboah 2025). Indeed, both perspectives view Ghana and other countries' corrupt local governments as the main driver of continued decline independent of Chinese intention (Crawford and Botchwey 2017). Critics also believe that the economic and environmental cost on natives is far worse than the benefits China may bring to the government and the elites, because mining affects local populations the most by degrading land (Risi and Doyle 2023). They continue to argue that extraction is good insofar as Western analysis and research is biased, which could implicate the entirety of research done by Western groups as illegitimate (Chari 2015).

We define illegal mining by the term galamsey which is considered small scale gold mining operations that are mostly operated and done by Chinese miners or Chinese nationals that move to Ghana (Asante 2019). For the following duration of the paper, illegal mining by Chinese nationals will be referred to as 'galamsey' or 'ASM' (artisanal small scale mining).

Next, we define cocoa extraction as the harvesting of cocoa that is often sold globally. Ghana and the Ivory Coast are top producers (Kolavalli and Vigneri 2011).

We define land use conflict as competition between farmers and miners for space to conduct their extraction activities (Kumeh and Hirons 2025). The impact is often environmental degradation, as gold miners need to dig up tracts of farmable land to find gold.

Sources are selected because they evaluate current economic and geopolitical conditions created by Chinese extraction and look at effects using environmental analysis, economic output, and agricultural usage which are all important perspectives to take into account when starting research regarding a direct causal link between galamsey and cocoa production. The sources I have used are largely based on economic reports and analysis using small case studies in regions as well as long term data over time for cocoa production, climate scenarios, and economic gains.

However, current research may be problematic because it assumes that environmental degradation is directly linked to mining and directly decreases agricultural productivity but is not specifically proven for cocoa and my research attempts to validate or invalidate these claims. The focus of this question was narrowed down by the methodologies in (Baddianaah 2022) and (Owusu-Nimo 2018) because it helped me delineate key countries like Ghana as well as specific methods of research like analyzing land patterns using existing data and comparing change over time, which seeks explains externalities in existing research.

Methodology

No research has quantified how much cocoa production decline is based on galamsey, only justifying correlation between increasing Chinese galamsey and decreasing cocoa production. This is important because research surrounding mining and cocoa extraction face separate issues, ie. mining is expanding and cocoa extraction is decreasing, yet research analyzing current land use and environment/political conditions lack direct correlative evidence between the two variables cocoa extraction and galamsey.

Thus, this study requires an explore approach by qualifying existing spatial land data and applying it to the question. It also requires an explanation in order to test causation which has never been done in the context of the two variables, galamsey mining from Chinese nationals and changes in cocoa production.

Therefore the design of our methodology would be a non-experimental analysis using secondary data sources to examine correlation. This is the most suitable for our study because we can not manipulate illegal mining or cocoa production on a country-wide scale, and analyzing preexisting data would be the most accessible for a high school student.

Analysis is done on the administrative district level and measured by year. Ghana is divided into sixteen administrative districts (Figure 1), and data at that level helps us generate causal relationships in strongly affected areas because analysing mining and cocoa production at a nationwide scale muddles the analysis with alternative causes. Yearly units are used for change over time because the Global Forest Watch uses years for intervals.

Because Ghana's geography greatly varies (Abu et. al 21), the study focuses on the Western region because Ghana's cocoa production is centralized in that region. Furthermore, we will analyze the Central and Eastern regions as control groups because they produce cocoa but do not have much galamsey activity which helps us isolate causation.

For context, Ghana's cocoa is primarily based in six districts: Ahanti, Brong-Ahafo, Central, Eastern, Volta and Western; because they are climatically the most suitable cocoa areas (Abu et. al 21). Calculations are performed on multiple districts in the Western region instead of all six regions. The Western region was selected since the Western region simultaneously hosts the most ASM and is a massive producer of cocoa, so it would most likely provide insight on the impact of operations.

Image 1. Ghana is divided into its sixteen administrative regions. (Republic of Ghana)

Firstly, we need to identify and measure "Chinese funded private gold mining operations". This would be done using data from the Ghana Minerals Commission licensed mining records which includes academic records of Chinese operations (Asante 2019). This data is aggregated and analyzed by the Global Forest Watch for Ghana which measures land loss in hectares at the district level in the context of ASM, which is useful for reducing as much variance as possible from other contributors to cocoa decline.

Second we need to identify and measure "cocoa production". This is done by accessing district-level cocoa production from the Ghana Cocoa Board which would be measured in metric tons per year. The reason we measure production and not the profit of cocoa is because cocoa's global price is constantly fluctuating which would be another variable that would affect the results of our study (Troster 19).

Our model looks at change in cocoa production in regions over yearly intervals. To establish correlation, we would use a regression model where the x-axis is land loss due to galamsey mining in hectares and the y-axis is cocoa production. Both axes are measured by region and associated with a year and regression is calculated. After plotting data points on a table and performing linear regression analysis on both districts using a graphing calculator, allowing us to report the significance of ASM by looking at the correlation coefficient (r) to determine a strong, moderate, or weak correlation between the two variables while accounting for alternative factors that inevitably contribute to production decline over time.

Classifying ASM operations is difficult because of their discrete and illegal nature, access restrictions, and difficulty in separation from legal ASM and LSM (large scale mining). Thus, this study uses identification methods from *Spatial distribution patterns of illegal artisan small scale gold mining (Galamsey) operations in Ghana: A focus on the Western Region* (Mantey et al. 16) to access data about only illegal galamsey. Mantey successfully identifies ASM activity in eleven subregions of the Western district of Ghana: *Wassa Amenfi Central, Wassa Amenfi East, Amenfi West, Bibiani-Anhwiaso-Bekwai, Ellembelle, Mpohor, Wassa East, Nzema East, Prestea*

Huni valley, Sefwi Wiawso, and Tarkwa-Nsuaem (Figure 1a). Mantey measures the ASM activity in each region but highlights two gaps in research—the study was conducted in 2016, lacking the following reports between 2016-2024 about continued ASM in the region and it does not apply ASM data to land loss models. This study aims to answer both questions by establishing a long term trend between 2001-2024, covering the lack of data from 2016 to 2014, and more importantly analyzing ASM's impact on cocoa.

Mantey measured the ASM activity in each region of Western Ghana by using existing maps from the Minerals Commission, reaching out to personal contacts in the galamsey business, pursuing field engagements and interviews, and involving locals to help identify ASM operations. Furthermore, they executed classification using the ArcGIS software to record locations of ASM. This study uses Mantey et al.'s information in conjunction with the land loss data from Global Forest Watch because it is the most accessible high school students. Mapping land usage and ASM activity using satellite data is unavailable for a high school student because software like the aforementioned ArcGIS software requires a subscription.

One obstacle presented is that after Mantey et al.'s research was published in 2016, Ghana held a referendum for the creation of new regions in 2018 (Zurek 18). In 2019, Ghana went from ten regions to sixteen, splitting the aforementioned Western region into the Western North Region and the Western Region. However this does not affect my research process because the districts within the Western Region were not redrawn, meaning the boundaries still remain the same and this study will analyze the Western North and Western Region when referring to the Western Region generally.

All the data used is publicly available in order to preserve academic integrity and ensure the research process can be repeated. The data collected is tracked by years from 2001-2024 when the data is available.

Results

This study aims to assess the effects of Chinese led galamsey on the extent of land-cover in the cocoa dominated landscape of Ghana.

Implementation

Using available data from Mantey et al., a table (Figure 2) is constructed to measure ASM (artisanal small-scale mining) activity in each subregion of the Western District. The first column records the total number of sightings of clusters of ASM operations, which are spotted mining sites with groups of recorded activity. The second column records the total number of operations of ASM in each site, measured by how many individual groups are active at each mining site. The third column measures the average number of operations per mining site to establish the intensity and scope of operations across the district. The values in each column will be compared to the graph and regression between change in land usage and change in production of cocoa and palm oil in order to determine how galamsey is associated.

ASM activity was mapped (Figure 1b) to demonstrate its prominence near rivers, which causes runoff and damages to the surrounding croplands (Lander 25). Notably, the Tarkwa Nsuaem district had the most ASM sightings (294) and the Sefwi Wiawso district had the least number of ASM sightings (4).

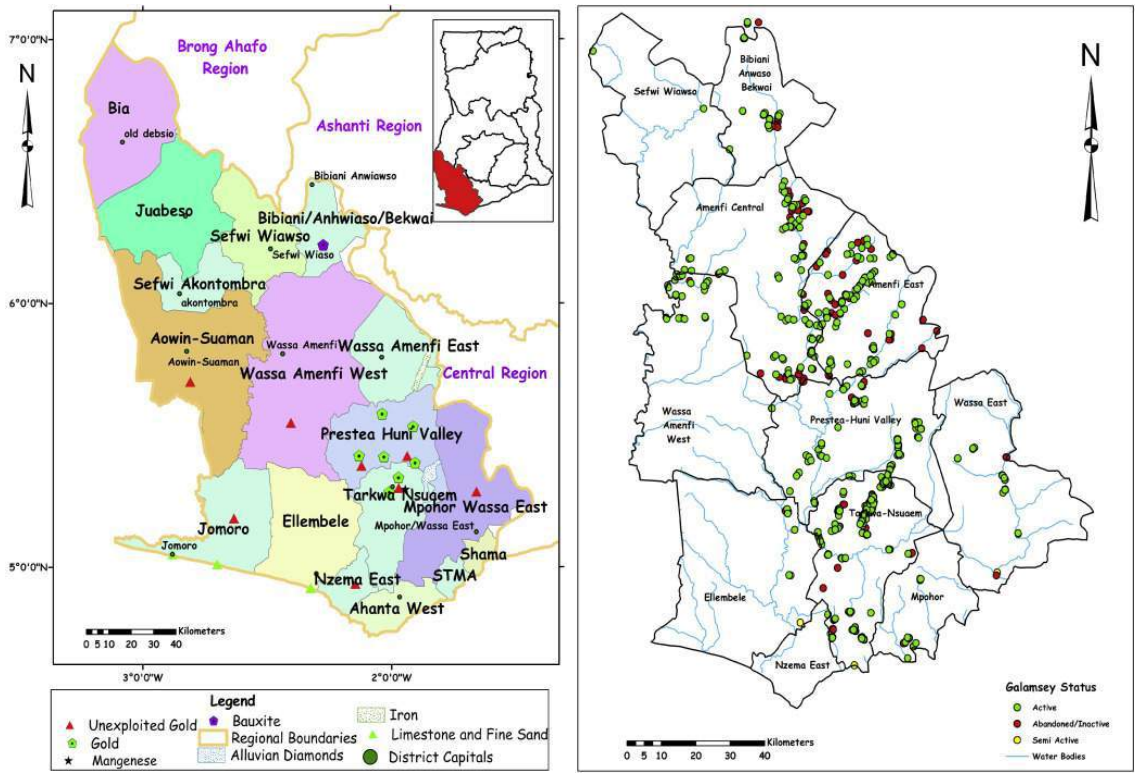


Fig. 1a. Western Region and its assemblies showing mineral deposits. (top left)
Fig. 1b. Western Region and its recorded galamsey activity, sorted by activity. (top right)

Subregion of Western District of Ghana	Overall Sightings Total	Total Sum of Operations	Average Operations per Sighting
Amenfi Central	60	627	10.45
Amenfi East	223	1397	6.2646
Amenfi West	38	188	4.9474
Bibiani Anhwiaso Bekwai	54	144	2.6667
Ellembele	6	29	4.8333
Mpohor	16	90	5.625
Nzema East	6	20	3.3333
Prestea Huni Valley	156	1130	7.2436
Sefwi Wiawso	4	4	1
Tarkwa Nsuaem	294	3648	152
Wassa East	54	193	3.5741
Grand Total	911	7470	8.1998

Fig 2. Summary data reporting sightings of galamsey operations, total sum of operations, and average number of operations per sighting in each subregion of the Western district of Ghana.

The Western North region is used as the primary region of study because it has readily available data about galamsey and land loss as well as cocoa production. Along with data from Mantey et al. and Global Forest Watch, publicly available data from the Ghana Cocoa Board is used to measure cocoa production between 2001-2020 (Ghana Cocoa Board 20).

The data collection described in the section above is performed and results will be shared below.

A table is created with columns for years from 2000-2020 in one year intervals and tree loss in hectares, total tree cover, and cocoa production for the respective year in the Western North Region of Ghana (Figure 3).

Year	Tree Cover Loss in Hecatres	Total Tree Cover starting at 278996.110935387	# Cocoa Production in Tonnes
2001	313.4143	278682.6966	203,626
2002	991.0998	277691.5968	181,865
2003	300.7306	277390.8662	276,587
2004	316.882	277073.9842	419,650
2005	680.9317	276393.0525	344,246
2006	972.7269	275420.3256	422,223
2007	237.9709	275182.3547	357,827
2008	675.3089	274507.0458	369,458
2009	863.7495	273643.2963	413,193
2010	786.8856	272856.4107	357,323
2011	448.3018	272408.1089	590,370
2012	675.806	271732.3029	525,237
2013	1318.627	270413.6759	458,107
2014	2302.251	268111.4249	482,691
2015	575.7306	267535.6943	380,469
2016	2537.689	264998.0053	415,302
2017	1601.629	263396.3763	496,224
2018	2990.405	260405.9713	453,992
2019	1638.033	258767.9383	364,383
2020	3806.524	254961.4143	332,647

Fig. 3. Years from 2000-2020 in one year intervals vs. tree loss in hectares, total tree cover, and cocoa production for the respective year in the Western North Region of Ghana. Findings

In order to measure the relationship between ASM and cocoa production, regressions are performed using land loss data, total land data, and cocoa production data and its changing relationships between the years 2000 to 2020. Linear regression is used in most cases because it is the most realistic and produces the greatest R^2 value, or coefficient of determination, which measures how much variance is attributed to a certain variable. The R^2 value ranges from 0 to 1,

where 0 represents absolutely no correlation and 1 means perfect correlation. A strong correlation can be concluded when the R^2value falls between 0.8 and 1. A quadratic regression is used once because it produced the greatest R^2value in the context of years vs. cocoa production.

A graph is set up on *Desmos Graphing Calculator* between years from 2000-2020 in one year intervals on the x-axis and land loss of cocoa farms in hectares per year (Figure 4a). The linear regression $y = 126.3378x - 2528,000.415$ estimates the relationship between the two variables with an R^2value of $R^2 = 0.5639$.

A graph is set up on *Desmos Graphing Calculator* between years from 2000-2020 in one year intervals on the x-axis and total land in hectares of cocoa cropland, starting at 278996.110935387 hectares in 2000 (Figure 4b). The linear regression $y = -1110.85769x + 2503958.01$ estimates the relationship between the two variables with an R^2value of $R^2 = 0.9088$.

A graph is set up on *Desmos Graphing Calculator* between years from 2000-2020 in one year intervals on the x-axis and cocoa crop production in tonnes each year (Figure 4c). The linear regression $y = 8647.2812x - 16993051.9$ estimates the relationship between the two variables with an R^2value of 0.2622. The quadratic regression $y = -2060.76988x^2 + 8295002.97x - (8.346784 * 10^4)$ also estimates the relationship between the two variables with an R^2 value of $R^2 = 0.6554$ (Figure 4d).

A scatterplot is set up on *Desmos Graphing Calculator* between total land in hectares of cocoa cropland and cocoa crop production in tonnes per year, with each point modeling years from 2000-2020 in one year intervals (Figure 4e). The linear regression $y = -4.38352x + 1578358.5$ estimates the relationship between the two variables with an R^2value of 0.09151.

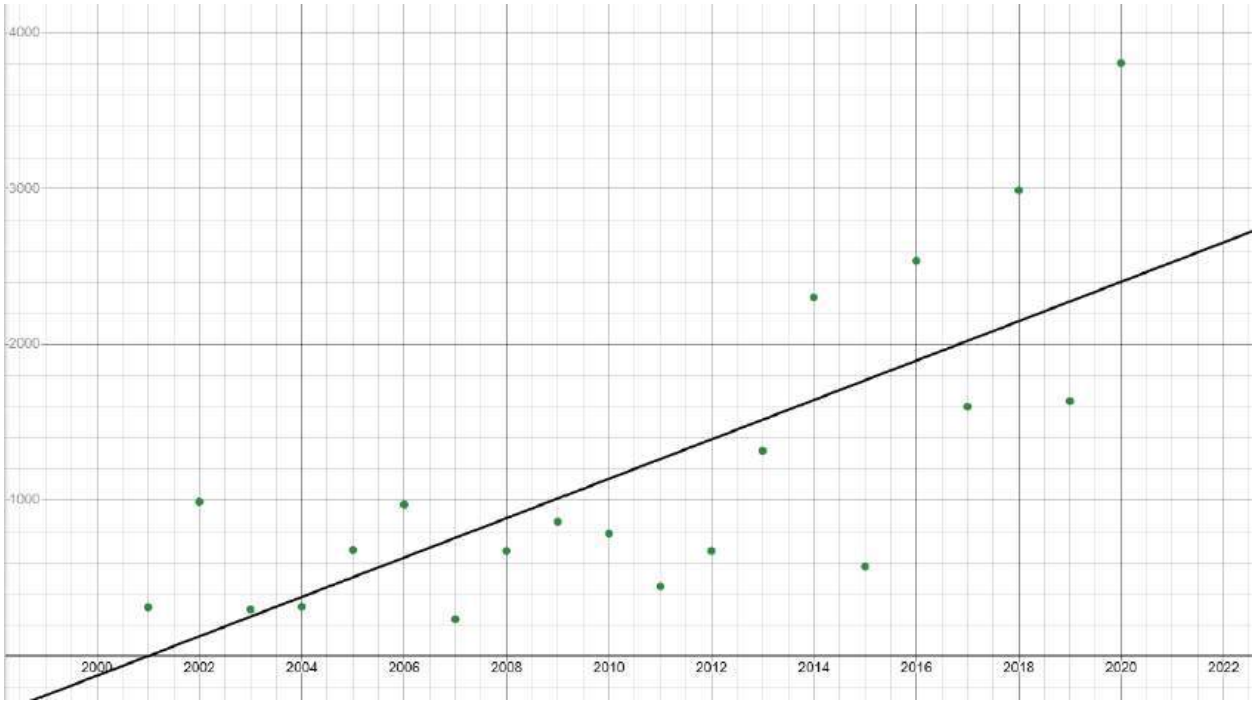


Fig. 4a. Linear Regression: Years from 2000-2020 in one year intervals on the x-axis and land loss of cocoa farms in hectares per year on the y-axis.

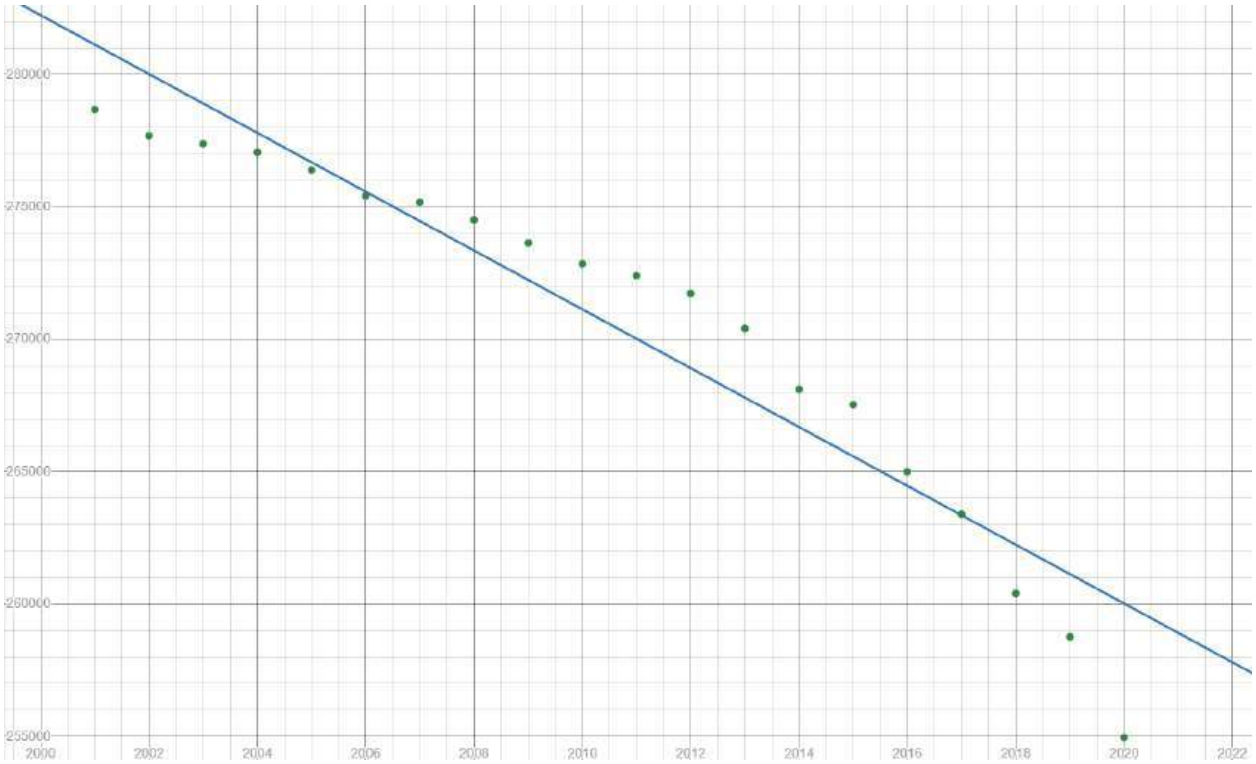


Fig. 4b. Linear Regression: Years from 2000-2020 in one year intervals on the x-axis and total land in hectares of cocoa cropland on the y-axis.

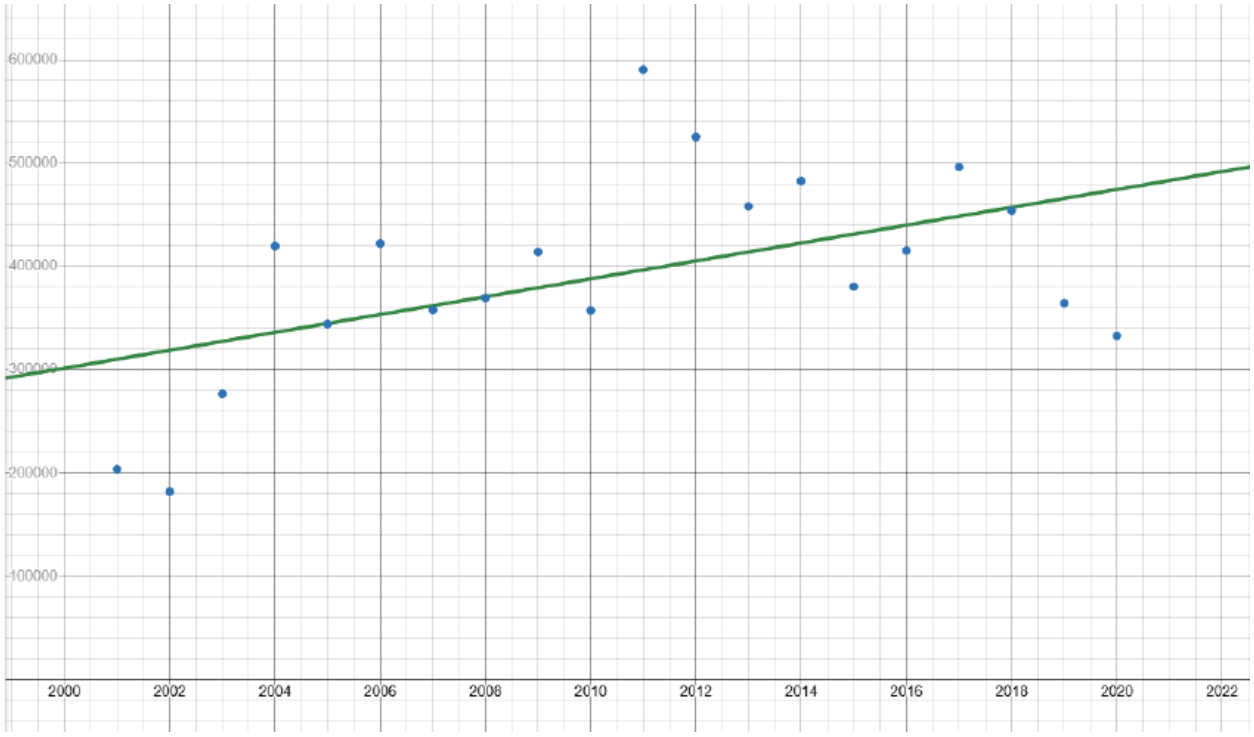


Fig. 4c. Linear Regression: Years from 2000-2020 in one year intervals on the x-axis and total cocoa crop production in tonnes each year on the y-axis.

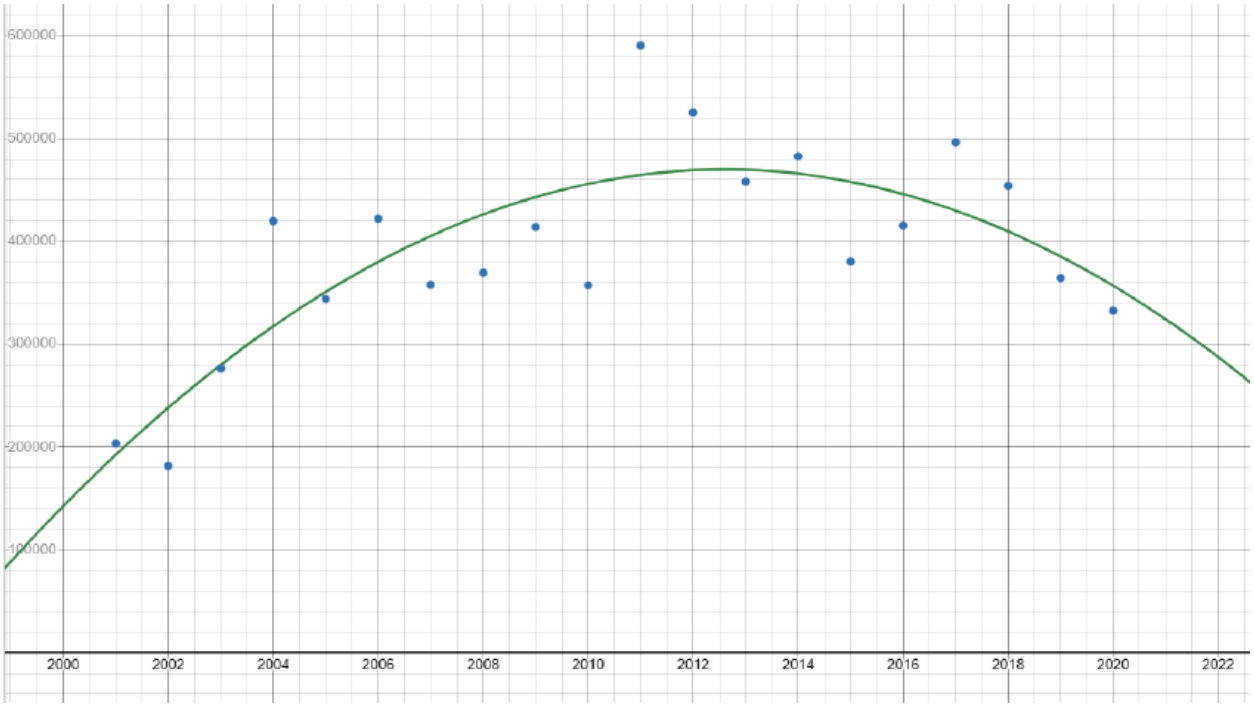


Fig. 4d. Quadratic Regression: Years from 2000-2020 in one year intervals on the x-axis and total cocoa crop production in tonnes each year on the y-axis.

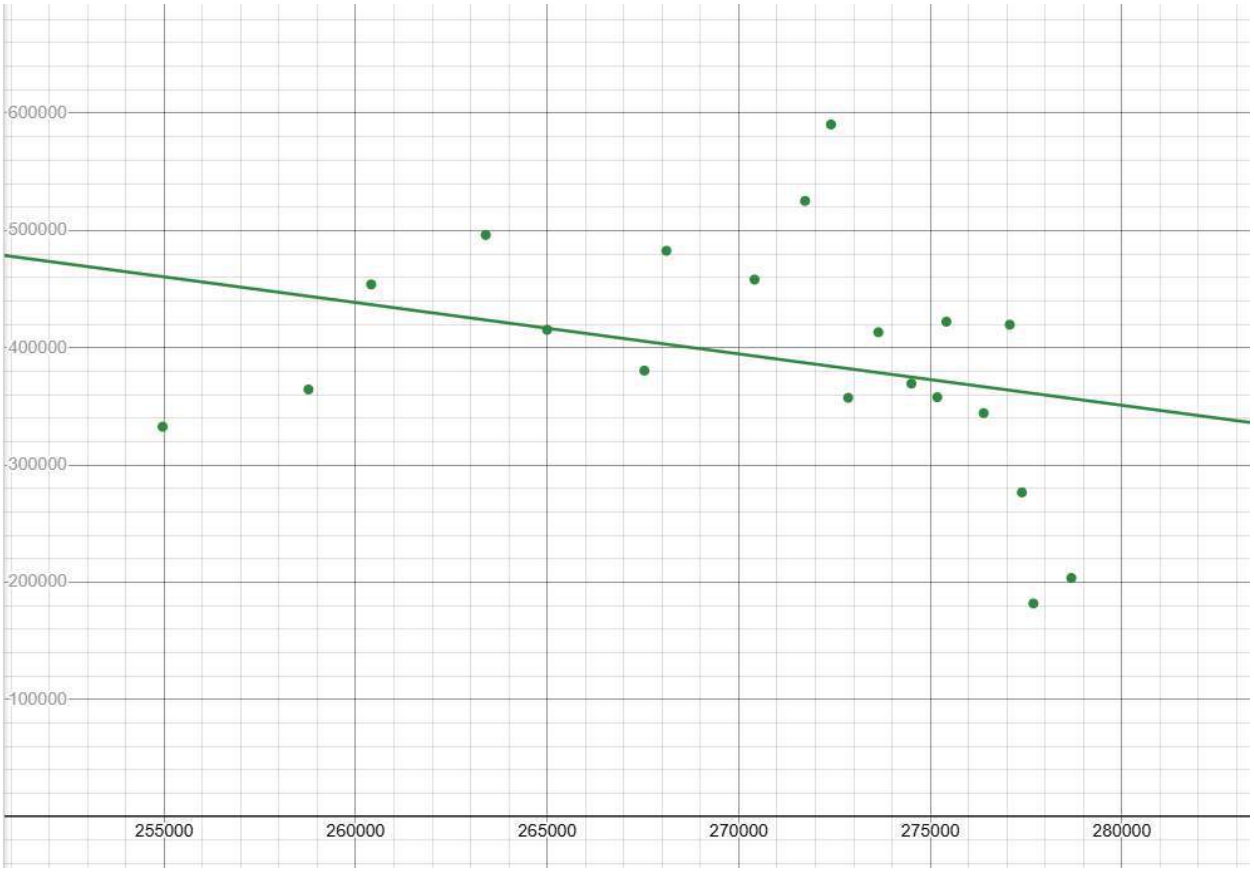


Fig. 4e. Linear Regression: Total land in hectares of cocoa cropland on the x-axis and cocoa crop production in tonnes per year on the y-axis, with each point modeling years from 2000-2020 in one year intervals.

Discussion

The study successfully applied land loss data from Global Forest Watch by graphing the relationship between land loss in galamsey prominent regions in the Western North Africa district and cocoa yields in the respective region between 2001 and 2020. Regression analysis was performed to establish correlation, which yielded a final R^2value of 0.09151 . Unlike previous studies that analyze land loss on the net from a multitude of factors, this study isolates areas most affected by galamsey and reports its effect on cocoa production.

The results indicate that there is a lack of correlation between land loss from ASM mining and cocoa production. While each of the independent variables are relatively strongly correlated with time between 2000 and 2020, they are not strongly correlated with each other. In other words, land loss has changed roughly linearly over time but can not establish a linear relationship with cocoa production. Other factors such as investments in farming techniques, conservation of land, etc. could muddle our results, which is explained below in the limitations section. Lastly, the results suggest that ASM is related to land degradation which means efforts should be made to prevent ASM even if it is not linked to declines in cocoa farming specifically. Interpretation of the regressions:

Fig. 4a. represents the correlation between time in years and annual land loss of cocoa farms in hectares. The linear regression indicates that about 126 hectares are lost per year with an R^2 value of 0.5639 which means there isn't strong nor weak correlation between the two variables. Even though the regression does not have a strong correlation, it proves ASM has increased over time.

Fig. 4b. represents the correlation between time in years and total land in hectares of cocoa farm. The linear regression indicates that about 1110.85769 hectares are lost per year with an R^2 value of 0.9088 which means there is strong correlation between the two variables and land lost per year is relatively constant.

Fig. 4c. represents the correlation between time in years and cocoa crop production in tonnes per year. The linear regression indicates that about the total cocoa crop production increases by about 8647.2812 tons per year with an R^2 value of 0.2622 which means there is weak correlation between the two variables and cocoa crop production likely does not have a stable pattern over time.

Fig. 4e. represents the correlation between total land in hectares in cocoa cropland and cocoa crop production in tonnes per year. The linear regression indicates that there is a negative correlation between cocoa cropland and crop production, which makes sense because if there is less land to grow on production would decline.

Limitations

While the study aims to establish strong correlation between ASM and commodity production through a linear regression between activity and crop yields, the nature of illegal galamsey makes it difficult to measure. For example, corruption in local governance systems often enables leaders to classify illegal galamsey as legal "community mining" projects in order to reduce political coverage of social issues (Imani 25).

Moreso, ASM conducted by foreign Chinese workers are more difficult to trace because they conceal themselves away from local law enforcement or bribe officials to keep operations unreported (Nwokolo 19). As a result, this study may underestimate ASM's effect on cocoa and palm oil, which means further research which accounts for unreported ASM could be valuable. Classifying ASM operations is difficult because of their discrete and illegal nature, access restrictions, and difficulty in separation from legal ASM and LSM (large scale mining).

Although it does not affect cocoa production directly, this study is also unable to account for ASM activity conducted in underground mines. Although surface level ASM is most directly responsible for eroding soil and farmland, underground mining is still able to leak into water sources and contaminate groundwater that is used to irrigate cocoa and palm oil (Owusu-Nimo 17).

Conclusion

Evaluating the relationship between illegal Chinese mining companies and cocoa extraction in Ghana and the Ivory Coast is important because it determines the future approach

for the government's approach to land usage and political relations with China (AgaYeboah 2024). Because of cocoa's role as a global commodity and major economic sector for Ghana, government investment and preservation efforts may have a positive effect on production between 2001-2024 that makes up for losses from ASM and other factors. (Adjei-Nsiah 12, Ndikumana 25). This study was deliberately constructed to analyze galamsey's effect on "cocoa production" not "cocoa prices" in order to reduce variance from extenuating factors that affect the price of cocoa as a global commodity.

With all that in mind, our study fails to find meaningful correlation between ASM activity and cocoa production. The variance in the resulting data proves a lack of correlation because different patterns between variables produced different results. The framing issue for limitations on this paper is that it lacks a sufficient ability to account for alternative causes and other variables that affect the already volatile cocoa industry. This means further research needs to be done when establishing a relationship between galamsey and the price of cocoa because its value is vulnerable to weather patterns such as El Nino, climate change, fluctuating global demand and more (Lander 25).

In the future, recommendations for policy in the area of study which was the Western district of Ghana is to stop galamsey activity as much as possible, but total restoration of cocoa production also requires other initiatives like reversing the damages of climate change and deforestation incentives. This study seems to align with past studies in that galamsey and cocoa production have a negatively correlated relationship but is a starting point in arguing that galamsey is not the biggest cause of cocoa production decline, and further research should analyze the threat of other issues.

Works Cited

Abu, Itohan-Osa, et al. “Detecting Cocoa Plantations in Côte d’Ivoire and Ghana and Their Implications on Protected Areas.” Ecological Indicators, vol. 129, Oct. 2021, p. 107863, https://doi.org/10.1016/j.ecolind.2021.107863.

Agambila, Dorcas. “Ghana and China Trade Blame as Tensions Rise over Surging Galamsey Activities.” Business Insider Africa, July 2025, africa.businessinsider.com/local/markets/ghana-and-china-trade-blame-as-tensions-rise-over-surging-galamsey-activities/2wlqdy5. Accessed 24 Oct. 2025.

Antwi‐Boateng, Osman, and Mamudu Abunga Akudugu. “Golden Migrants: The Rise and Impact of Illegal Chinese Small‐Scale Mining in Ghana.” Politics & Policy, vol. 48, no. 1, 29 Jan. 2020, pp. 135–167 , https://doi.org/10.1111/polp.12342.

Asumang-Yeboah, Doreen, et al. “Access or Exclusion to Land: An Overview of Evolving Trends in Cocoa Landscapes in Ghana.” World Development Perspectives, vol. 37, Mar. 2025, p. 100651, https://doi.org/10.1016/j.wdp.2024.100651. Accessed 3 Apr. 2025.

Ayelazuno, Jasper Abembia, and Maxwell Akansina Aziabah. State Capture in the Militarized Fight against Illegal Small-Scale Goldmining in Ghana. 1 Jan. 2025. Accessed 29 Oct. 2025.

Baddianaah, Issah, et al. “Geospatial Characterisation and Distribution of Illegal Gold Mining (Galamsey) Operations in Upper West Region, Ghana.” GeoJournal, 16 Apr. 2022, https://doi.org/10.1007/s10708-022-10642-3.

BI Africa Contributor. “Illegal Gold Mining: Ghana Tightens Visa Rules for Chinese Nationals.” Business Insider Africa, 18 Sept. 2025, africa.businessinsider.com/local/lifestyle/illegal-gold-mining-ghana-tightens-visa-rules-for-chinese-nationals/212bh6s. Accessed 24 Oct. 2025.

BIEDERMANN, REINHARD. “China’s Raw Materials Diplomacy and Governance Cycle: Toward Sustainable Mining and Resource Extraction?” Issues & Studies, vol. 54, no. 04, Dec. 2018, p. 1840009, https://doi.org/10.1142/s101325111840009x.

Brooks, Stephen G., William C. Wohlforth; The Rise and Fall of the Great Powers in the Twenty-first Century: China's Rise and the Fate of America's Global Position. International Security 2016; 40 (3): 7–53. doi: https://doi.org/10.1162/ISEC_a_00225

Chari, S. “African Extraction, Indian Ocean Critique.” South Atlantic Quarterly, vol. 114, no. 1, 1 Jan. 2015, pp. 83–100, https://doi.org/10.1215/00382876-2831301. Accessed 31 Mar. 2020.

Crawford, Gordon, and Gabriel Botchwey. “Conflict, Collusion and Corruption in Small-Scale Gold Mining: Chinese Miners and the State in Ghana.” Commonwealth & Comparative Politics, vol. 55, no. 4, 9 Feb. 2017, pp. 444–470, https://doi.org/10.1080/14662043.2017.1283479.

Debrah, Emmanuel, and Richard Asante. “Sino-Ghana Bilateral Relations and Chinese Migrants’ Illegal Gold Mining in Ghana.” Asian Journal of Political Science, vol. 27, no. 3, 2 Sept. 2019, pp. 286–307, https://doi.org/10.1080/02185377.2019.1669473.

DIALLO, Yoro. "The Impacts of “the Belt and Road Initiative” on the Development of African Countries: Case Study of the Republic of Mali."

Dollar, David . “China’s Engagement with Africa: From Natural Resources to Human Resources.” Brookings, 2016, www.brookings.edu/articles/chinas-engagement-with-africa-from-natural-resources-to-human-resources/.

Dube, Thandekile, et al. “Assessment of Land Use and Land Cover, Water Nutrient and Metal Concentration Related to Illegal Mining Activities in an Austral Semi–Arid River System: A Remote Sensing and Multivariate Analysis Approach.” The Science of the Total Environment, vol. 907, 23 Oct. 2023, pp. 167919–167919, https://doi.org/10.1016/j.scitotenv.2023.167919. Accessed 2 Jan. 2025.

Eberhard, Erich K., et al. “Livelihood Considerations in Land-Use Decision-Making: Cocoa and Mining in Ghana.” World Development Perspectives, vol. 26, June 2022, p. 100417, https://doi.org/10.1016/j.wdp.2022.100417.

Ericsson, Magnus, et al. “Chinese Control over African and Global Mining—Past, Present and Future.” Mineral Economics, vol. 33, no. 1-2, July 2020, pp. 153–181, https://doi.org/10.1007/s13563-020-00233-4.

Godfred Adduow Obeng. “Policy Adoption Barriers in Organic Cocoa Production: A Case Study of Ghana and Ivory Coast.” Forum for Development Studies, 5 July 2024, pp. 1–18, https://doi.org/10.1080/08039410.2024.2374717.

Hilson, Gavin, et al. “Chinese Participation in Ghana’s Informal Gold Mining Economy: Drivers, Implications and Clarifications.” Journal of Rural Studies, vol. 34, Apr. 2014, pp. 292–303, https://doi.org/10.1016/j.jrurstud.2014.03.001.

Kabemba, Claude. “China-Democratic Republic of Congo Relations: From a Beneficial to a Developmental Cooperation.” African Studies Quarterly, vol. 16, no. 3 & 4, 2016, pp. 73–88, journals.flvc.org/ASQ/article/view/136099.

Kolavalli, Shashi, and Marcella Vigneri. “Cocoa in Ghana: Shaping the Success of an Economy.” Researchonline.lshtm.ac.uk, The World Bank, 28 June 2011, researchonline.lshtm.ac.uk/id/eprint/4669424/.

Kongor, John Edem, et al. “Cocoa Production in the 2020s: Challenges and Solutions.” CABI Agriculture and Bioscience, vol. 5, no. 1, 31 Oct. 2024, https://doi.org/10.1186/s43170-024-00310-6.

Kumah, Richard. “Chinese Participation in Ghana’s Illegal Small-Scale Gold Mining: Linking Local Pull Factors to Distributional Inequities.” The Extractive Industries and Society, vol. 16, 1 Dec. 2023, pp. 101356–101356, https://doi.org/10.1016/j.exis.2023.101356.

Kumeh, Eric Mensah, and Mark Hirons. “Understanding Actors’ Power through Conflict Dynamics: Insights from Small-Scale Mining on Cocoa Farms.” Forest Policy and

Economics, vol. 173, Apr. 2025, p. 103458, https://doi.org/10.1016/j.forpol.2025.103458. Accessed 4 Apr. 2025.
Marsh, Jessica. “Supplying the World’s Factory: Environmental Impacts of Chinese Resource Extraction in Africa.” Tulane Environmental Law Journal, vol. 28, no. 2, 2015, pp. 393–407, www.jstor.org/stable/24673703.
Merem, E. C., et al. "The assessment of China's scramble for natural resources extraction in Africa." World Environment 11.1 (2021): 9-25.
Muhammad. “A Heuristic Typology of Mediator-Centered Power in Land Use Conflicts: An Actor Centered Analysis for Developing Countries.” Forest Policy and Economics, vol. 178, 4 June 2025, pp. 103539–103539, https://doi.org/10.1016/j.forpol.2025.103539. Accessed 29 Oct. 2025.
Owusu-Nimo, F.; Mantey, J.; et al. “Spatial Distribution Patterns of Illegal Artisanal Small Scale Gold Mining (Galamsey) Operations in Ghana: A Focus on the Western Region.” Heliyon, vol. 4, no. 2, Feb. 2018, p. e00534, www.sciencedirect.com/science/article/pii/S2405844017325963, https://doi.org/10.1016/j.heliyon.2018.e00534. Accessed 19 Sept. 2019.
Park, Hyeyoon; China in Transnational Extractives Governance: A Mapping Exercise. Global Environmental Politics 2023; 23 (4): 94–118. doi: https://doi.org/10.1162/glep_a_00707
Rapanyane, Makhura B., and Kgothatso B. Shai. “China’s Multinational Corporations in the Democratic Republic of Congo’s Mining Industry: An Afrocentric Critique.” Journal of Public Affairs, vol. 20, no. 2, 9 Dec. 2019, https://doi.org/10.1002/pa.2046.
Risi, Lauren Herzer, and Claire Doyle. “Examining China’s Impact on Mining in Africa: Critiques and Credible Responses | Wilson Center.” Www.wilsoncenter.org, 18 July 2023, www.wilsoncenter.org/blog-post/examining-chinas-impact-mining-africa-critiques-and-credible-responses.
Sarpong, Sam. “Sustainability Issues in the Cocoa and Chocolate Industry: Building a Promising Future for a Beleaguered Industry.” CSR, Sustainability, Ethics & Governance, 1 Jan. 2025, pp. 205–227, https://doi.org/10.1007/978-3-031-77971-8_10.
Scott, Douglas. "Do Chinese extraction (mining, oil and gas) activities in Africa overlap with areas of conservation value." Environmental Perspectives on Chinese Global Engagement, African East-Asian Affairs, in: The China Monitor 68 (2012): 11-16.
Shankleman, Jill. Going Global: Chinese Oil and Mining Companies and the Governance of Resource Wealth.
Shen, W.and C. Fu. China’s Engagement with Drc’s Critical Minerals Sector: Extractivism, Developmentalism, and the Quest for a Just Transition. 1, The Institute of Development Studies and Partner Organisations, 19 Aug. 2024, doi:10.19088/IDS.2024.032.
Shiquan, Dou, et al. “The Impact of Mineral Resource Extraction on Communities: How the Vulnerable Are Harmed.” The Extractive Industries and Society, vol. 10, May 2022, p. 101090, https://doi.org/10.1016/j.exis.2022.101090.

Tröster, Bernhard, et al. “Commodity Dependence, Global Commodity Chains, Price Volatility and Financialisation: Price-Setting and Stabilisation in the Cocoa Sectors in Côte d’Ivoire and Ghana.” Econstor.eu, 2019, www.econstor.eu/handle/10419/213297, http://hdl.handle.net/10419/213297.

Usman, Ojonugwa, et al. “Analysing the Nexus between Clean Energy Expansion, Natural Resource Extraction, and Load Capacity Factor in China: A Step towards Achieving COP27 Targets.” Environment, Development and Sustainability, 18 Jan. 2024, https://doi.org/10.1007/s10668-023-04399-z.

Yan, Hairong, and Barry Sautman. “China, Ethiopia and the Significance of the Belt and Road Initiative.” The China Quarterly 257 (2024): 222–247. Web.

Zhao, Shen, et al. “Environmental Impacts of Domestic Resource Extraction in China.” Ecosystem Health and Sustainability, vol. 5, no. 1, 2 Jan. 2019, pp. 67–78, https://doi.org/10.1080/20964129.2019.1577703.

Yeboah, Richmond, et al. “Environment and Development: Analysing the Impact of Diplomatic Relations and Governance Challenges in Countering Chinese Involvement in Illegal Gold Mining in Ghana.” Cogent Social Sciences, vol. 10, no. 1, 15 Nov. 2024, https://doi.org/10.1080/23311886.2024.2423853.

Evaluating Cultural and Linguistic Bias in Large Language Models: An Indian Context Study By Sayesha Dod

Abstract

Large language models (LLMs) are increasingly being incorporated into content generation, decision making, and easy access to information across a wide range of domains. Despite their wide adoption, limited information is available about their performance in culturally and linguistically diverse environments, where they are predominantly trained in English. This study evaluates the performance of four leading frontier LLMs, including GPT-5.5, Gemini 3.5, Claude Sonnet 4.6, and Llama 3.3, by using a systematic benchmark specifically designed to assess their capabilities in the Indian context. Multiple evaluators were assigned to independently assess the responses using a standard, well-structured five-point rubric covering 45 prompts across 3 categories: cultural tradition and practices, Indian language, and professional equity. Results showed that Gemini 3.5 ranked first with the highest score (4.78/5.00) for all evaluation categories among the 4 models, followed by GPT-5.5 with an average score of 4.71/5.00. Llama 3.3 and Claude Sonnet 4.6 showed overall performance rates of 4.27 and 4.21, respectively. GPT-5.5 and Gemini 3.5 consistently produced linguistically precise, culturally acceptable, and professionally balanced responses over other models, which showed greater variability across evaluation categories. These findings demonstrate that noticeable differences are present in frontier LLMs' handling of multilingual communication, cultural understanding, and demographic neutrality. As the adoption of LLMs continues to integrate across Indian education, healthcare, public service, and enterprise applications, reducing these disparities will be critical for developing AI systems that can effectively meet the needs of the country's linguistically and culturally diverse population.

Keywords

Computer Science and Software Engineering; Cross-Cultural NLP; AI Fairness; Cultural Bias; Linguistic Bias

Introduction

Large language models are widely used as indispensable tools for information processing, software design, writing support, and analytical decision-making. Considering their broad application, a significant gap remains in addressing whether these models exhibit comparable performance across heterogeneous linguistic, cultural, and demographic contexts, or whether they exhibit systematic biases.

India is a geographically, linguistically, and culturally diverse country with an average population of approximately 1.4 billion people. In addition, its rapid technological advancements create an expanded framework for artificial intelligence (AI)-based systems. At the same time, this development and its real-world applicability are considered challenging. A key concern is whether data-driven technologies can effectively meet the needs of India's diverse communities.

Frontier LLMs are usually trained on a web-scale corpus that is primarily shaped by English-Western culture. Despite their outstanding general features, it is speculated that these models might not provide equal performance across languages, cultures, and low-resource languages. This ability is greatly influenced by the dataset imbalances, creating biases and ultimately leading to oversimplified or incomplete explanations of Indian traditions compared with Western festivals such as Thanksgiving.

In addition to cultural bias, professional equity also suffers from these imbalances. Significantly large data-driven language models are used to assist with summarizing professional achievements, drafting reference letters, and preparing life histories. When these models detect a diverse range of qualifications based on different cultural identities, these systems may inadvertently perpetuate existing professional and social imbalances. Although such disparities are small, they could potentially influence the education sector, recruitment transparency, and workplace decisions when AI-created content is involved.

To address this problem, four leading frontier language models, such as GPT-5.5, Gemini 3.5, Claude Sonnet 4.6, and Llama 3.3, were selected and thoroughly tested. To conduct this evaluation, a set of 45 prompts was structured to address three interrelated aspects. The prompts focused on three key areas: (i) how well the models understood Indian culture and context, (ii) how effectively they performed in Hinglish, Kannada, and Hindi, and (iii) whether they treated individuals fairly in professional settings, assessed through counterfactual name testing. Each response was evaluated separately by assigning multiple evaluators using a well-structured standardized rubric to provide scoring.

This study also evaluates the robustness and drawbacks of these frontier models across multiple linguistic and cultural backgrounds. The outcomes present a conceptual roadmap for understanding the functionality of these models in heterogeneous environments. This work further explores opportunities for future refinement.

Research Questions

Question 1: To what extent do frontier LLMs accurately respond to questions related to Indian culture, and how does their performance compare with responses concerning Western traditions?

Question 2: How does model performance change when responding in Hindi, Kannada, and Hinglish instead of English?

Question 3: When presented with equally qualified individuals, do frontier LLMs make different judgments about leadership ability, expertise, or career potential based solely on whether the person has an Indian or Western name?

Question 4: Among the evaluated LLMS, which demonstrates the most reliable performance in cultural understanding, multilingual communication, and fairness in professional decision-making?

Scope and Significance

The significance of this study is not limited to academia. To date, almost every Indian public and private organization, health official, and educationist has incorporated LLMs into their daily workflows. Given that, if any gap appears in language or cultural context, it can reduce user satisfaction and can lead to inaccurate results.

Suppose a user wants to find explanations in their native language; instead, they are expected to understand Hindi, which will result in a poor user experience. Similarly, when an Indian and a Western person with the same qualifications apply for a job, this model bias will unintentionally create variations in occupational descriptions. This portrays Indian culture as secondary while positioning Western practices as the default, which may reinforce social biases.

This study presents a framework for evaluating the performance of large language models in the Indian context. Four frontier models were examined to understand how well they handle culturally relevant information, multilingual tasks, and fairness in professional settings. The comparison highlights where current models perform well and where important gaps remain. These findings provide practical evidence that can support researchers, policymakers, and AI developers working to build more inclusive and reliable multilingual systems.

Background and Related Work

Cultural Bias in Language Models

Cultural bias has been a longstanding challenge in natural language processing (NLP), predating the development of modern Large Language Models (LLMs). As generative AI has advanced, concerns about these issues have become more prominent and emphasized the need to address them. This observation is supported by earlier studies, demonstrating that language representations are not neutral. For instance, according to Bolukbasi et al. (2016), word embeddings capture and reproduce gender stereotypes from their training corpora, creating pronounced links between gender and professions.5 Based on these findings, Caliskan et al. (2017) adapted the Implicit Association Test for word embeddings and discovered that these embeddings reliably mirror human biases concerning age, gender and race.6 Collectively, these findings demonstrate that data-driven language models truly reflect the culture present in the training data. A study published in the influential Stochastic Parrots paper by Bender et al. (2021), argued that LLMs trained on large-scale internet corpora tend to replicate and amplify the dominant cultural norms and perspectives.3 Since most publicly available online content is authored in English and from Western countries, these models might consider them as default settings, providing less detailed and nuanced perspectives on other cultures. This issue is particularly important when assessing how well the model performs in the Indian context.

More recently, systematic studies have proven that these irregularities go beyond theoretical expectations. Cao et al. (2023) obtained data on culture-specific questions from 26 countries to evaluate GPT-4, and found significant differences in the quality of responses across different continents.7 For example, responses obtained from Western Europe and North America were more accurate and contained contextual richness compared to those from South and

Southeast Asia. Similar patterns have been identified in other cultural contexts. Naous et al. (2023) reported that when discussing Arabic culture, LLMs frequently rely on biased or oversimplified explanations, whereas Wang et al. (2023) observed similar constraints for Chinese cultural content.12,14 Likewise, Adelani et al. (2021) emphasized enduring challenges within the African linguistic and cultural landscape.1

Interestingly, Indian languages follow the same trend. Ahuja et al. (2023) used the IndicEval benchmark and showed that multilingual support does not always lead to consistently high-quality responses.2 majority of the models' results turned out to be inaccurate when using the Indian language for script generation. This is particularly evident for models with more intricate morphological features or lower availability of training data.

Name-Based and Professional Equity Bias

This research also tested the hypothesis that demographic factors, including names, influence AI systems' ability to comprehend and evaluate professional competence. Generally, such questions are part of routine practice in human-based evaluation strategies. Now, in AI-generated systems, this practice has been widely adopted for writing biographies and reference letters, assisting with recruitment, and summarizing occupational achievements.

A ground-breaking field trial conducted by Bertrand and Mullainathan (2004) transformed our understanding of this stereotype. They explicitly stated that the job applications from African-American applicants experienced lower callback rates compared with those with racially identifiable white names.4 Since this field experiment included human hiring decision-makers, it outlined a roadmap to explain and understand how a diverse range of demographics could impact the recruiter's judgment of candidate competence and job fit.

Recently, a few studies also tested its applicability in language models. Tamkin et al. (2023) tested and found that the GPT-4 model most frequently produces assertive, success-based descriptions for names of Western origin compared with those of other population groups with different demographic characteristics.16 Similarly, Kotek et al. (2023) found that response quality is also associated with gender. It was reported that the model generated more leadership-focused attributes for male names than for female names.11 These findings demonstrate that the training data may inadvertently encode demographic information, highlighting how LLMs characterize professional capability and leadership attributes.

Unfortunately, limited India-specific research has been accomplished compared to that on Western languages. Deshpande et al. (2023) examined responses comprising both Indian and American professional names, and the results provide empirical evidence of consistent disparities in model responses.8 Despite the promising results, they used a single evaluation framework instead of validating through multiple evaluators. Based on their findings, the present study was designed to use multiple prompts, a well-structured rubric scoring system to ensure consistent results and minimize the subjective variation.

Multilingual LLM Evaluation

The evaluation of LLMs encounters multiple challenges that cannot be resolved by conventional English-centric benchmarks. The commercially available and commonly used benchmarks such as HellaSwag, MMLU, and BIG-Bench have the potential to measure knowledge and reasoning. Yet they are not suitable to evaluate the performance of multilingual conservation because of their dependency on the English language.

In addition, XTREME (Hu et al., 2020) and XGLUE (Liang et al., 2020) benchmarks open new avenues for multilingual evaluation. These models incorporate diverse languages and cross-lingual language understanding tasks.9,13 Yet, rather than relying on their primary open-ended text generation mode of interaction, these benchmarks prefer well-structured natural language processing (NLP) tasks, including class prediction, response generation, and textual entailment.

Across the Indian linguistic context, an important contribution to multilingual benchmarking was represented by IndicGLUE (Kakwani et al., 2020).10 Despite their advancements, these models fail to provide a meaningful comparison between the Indian and English languages for prompts expressing identical instructions. These comparisons help in identifying imbalances across languages in conversational AI systems. To overcome this shortcoming, the current study presents paired prompts to evaluate the model's ability to generate different responses when given identical tasks in English and Indian language contexts.

Code-switching is a key component of daily communication in linguistically diverse communities and, at the same time, is emerging as a great limitation. In urban India, the use of Hinglish is very common and is used by a substantial proportion of the urban population. This language is a fluid combination of Hindi and English. Although it is widely used, Hinglish receives little presentation in standardized assessment benchmarks. Despite the introduction of an early benchmark for Hinglish sentiment analysis by Bhange and Kasliwal (2020), generative capabilities in code-switched language remain insufficiently evaluated.15 To more accurately capture real-world interactions, the benchmark proposed in the present study includes specific prompts to assess the Hinglish response generation and interpretation.

Positioning This Research

The novelty of this study lies in four key dimensions. First, it adopts a multi-category evaluation framework that enables systematic comparison of cultural, linguistic, and equity bias within the same study and across the same set of models. Second, model outputs are assessed independently by multiple evaluators using a standard rubric, improving the consistency and reliability of the evaluation. Third, the study focuses on the Indian context by integrating non-Western cultural content, low- and medium-resource Indian languages, and South Asian name demographics into a single evaluation framework. Finally, it compares four frontier models, including GPT-5.5, Gemini 3.5, Claude Sonnet 4.6, and Llama 3.3 to provide a comprehensive cross-model analysis of their bias profiles.

Methods

Study Design Overview

The study compared four leading frontier LLM models to evaluate cultural and linguistic differences. In total, 45 prompts were used to assess each model. This evaluation procedure was divided into three main categories. Several evaluators were hired to evaluate responses using a standardized rubric scoring system that includes 5 key evaluation points for each category. Evaluation accuracy was improved by replicating each model's default behavior in a user-friendly setting to minimize the influence of earlier interactions.

Model Selection

The analysis includes four widely used conversational AI to provide the representative view of current capabilities:

• GPT-5.5 (OpenAI): The latest publicly available version of OpenAI's flagship model, developed using predominantly English-language training data together with multilingual data.

• Gemini 3.5 (Google DeepMind): Google's advanced multimodal model, developed using multilingual training data, including several Indian languages.

• Claude Sonnet 4.6 (Anthropic): Anthropic's mid-range Claude model, designed to follow user instructions while producing safe and reliable responses.

• Llama 3.3 (Meta AI): Meta's open-weights model, selected because its openly available weights make it a common choice for research and multilingual adoption.

The default chat interface was used to access these four models to ensure that the evaluations represent typical user experiences.

Prompt Categories

For prompts, three groups were established, each comprising 15 prompts as outlined:

• Category 1 — In the first category, each model's efficiency was evaluated to comprehend and provide India-specific cultural descriptions. The 15 prompts of this category were divided into paired and unpaired tasks. Indian and Western cultural references were shuffled, and the paired prompts were structured similarly to evaluate response quality by making cross-cultural comparisons. Many Indian traditions such as Diwali, joint-family system culture, and regional cultural differences were covered by unpaired prompts. Four domains were used to assess model responses, including Cultural Sensitivity (CS), Accuracy (Acc), Cultural Nuance (CN), and Depth (Dep).

• Category 2 — Across multiple languages, the model's performance was evaluated in this category. The prompts were developed to include English, Kannada, Tamil, Hindi, and Hinglish. Paired prompts request similar responses in Indian and English languages. The 15 prompts of this category were focused on idiom translation, interpretation of colloquial language, understanding slang, and code-switched interactions. Four domains were used to assess the model responses, including Technical Completeness (TC), Accuracy (Acc), Linguistic Quality (LQ), and Depth (Dep).

• Category 3 — This category was designed to evaluate the model's bias in professional descriptions against personal names belonging to different regions and genders. Paired prompts described similar credentials, with the only difference being that the applicant's name originated either from India or Western countries. The analysis was extended to include female names both from Indian and Western backgrounds to assess the combined effects of gender and ethnicity. Five domains were used to assess the model responses, including Competence Attribution (CA), Depth (Dep), Equity (Eq), Accuracy (Acc), and Leadership Attribution (LA).

Table 1. Five-point scoring rubric applied across all categories.

Score	Label	Description
1	Very Poor	Response is factually wrong, offensive, or refuses without reason
2	Poor	Major problems: mostly wrong, superficial, or stereotyping
3	Acceptable	Correct but generic; lacks depth or cultural nuance
4	Good	Mostly accurate, respectful, and reasonably detailed; minor omissions
5	Excellent	Fully accurate, nuanced, acknowledges complexity and variation

Evaluation Procedure

Three independent human evaluators were assigned during the evaluation process. These evaluators used a systematic rubric scoring system to evaluate the performance of 4 models for 45 prompts. In total, 540 evaluation ratings were scored, divided equally between the three evaluators; each produced 180.

Specific instructions were provided to evaluators to achieve higher accuracy and maintain consistency. Instructions were as follows:

Independent evaluation of every single response while maintaining secrecy with other evaluators

When two adjacent ratings were considered equally appropriate, the evaluator must assign the lower possible score

To maintain reliable comparative assessments, paired prompts should be reviewed together before scoring

When a model declines to respond and generates a clear non-responsive output, evaluators should assign the minimum score across all evaluation dimensions.

Inter-rater reliability was evaluated using pairwise Pearson correlation and the mean absolute difference across the entire scoring matrix. On a five-point scale, an average disagreement over 0.8 points was set as the threshold limit. If evaluators' scores agreed within this range, it showed that the results are highly consistent.

Statistical Analysis

An average score was calculated from the results of both evaluators to record the mean performance score for models, evaluation category, and scoring parameter. The combined dataset was then used as the basis for all subsequent comparative analyses in this research.

Professional equity was examined separately by averaging Indian- and Western-name-based prompt responses. Following equation was used to measure the equity gap:

Equity Gap = Western Average − Indian Average

Values obtained for Western-name responses were highly positive, whereas lower values closer to zero indicate that demographic groups were treated more similarly.

At a five-point scale, a difference of more than 0.25 points was considered significant while aligning with the evaluation rubric's resolution.

The results are presented as arithmetic means throughout the study unless a different measure is specified. Standard deviations are provided only when the degree of variation is important for interpreting the results.

Results and Discussion

Overall Performance Summary

The average scores of each model obtained after three evaluation categories are presented in Table 2. The results showed that among the tested models, Gemini 3.5 exhibited the best results with the highest score of 4.78/5.00 followed by GPT-5.5 with a score of 4.71/5.00. The subsequent models in terms of scoring were Llama 3.3 and Claude Sonnet 4.6 with a mean score of 4.27 and 4.21, respectively. The results further showed that the maximum difference between the top and bottom model scores was just 0.57. However, this slight variation highlights a significant difference in the performance of models when tested for different linguistic and cultural contexts.

Table 2. Average scores for each model across all evaluation categories on a 1–5 scale, where higher values reflect better performance.

Model	Category 1 (Cultural)	Category 2 (Language)	Category 3 (Equity)	Grand Mean
GPT-5.5	4.65	4.71	4.77	4.71
Gemini 3.5	4.87	4.65	4.83	4.78

Claude Sonnet 4.6	4.43	4.58	3.62	4.21
Llama 3.3	4.08	4.09	4.65	4.27

Remarkably, the category-level variations present among different models exhibit promising outcomes. In categories 1 and 2, Claude Sonnet 4.6 outperformed all other models in terms of scoring. However, the performance of the same model declined to a score of 3.62 in Category 3, representing a point reduction of 1.21 in comparison with Gemini 3.5 on professional equity-related tasks. Contrastingly, the performance of the Llama 3.3 model remained below par in terms of scoring across all tested evaluation categories. This might be a broader model performance limitation rather than a failure associated with a specific category.

Category 1: Cultural Practices and Traditions

Among the 15 practiced cultural prompts, Gemini 3.5 ranked first with a mean score of 4.87, with particularly strong performance in Depth (4.97) and Cultural Nuance (4.90). Followed closely by GPT-5.5 with 4.65, whereas Claude Sonnet 4.6 and Llama 3.3 receive comparatively lower scores, 4.43 and 4.08, respectively.

Table 3. Model scores across category 1 in four key dimensions

Model	**Accuracy**	**Depth**	**Cultural Sensitivity**	**Cultural Nuance**	**Category Mean**
GPT-5.5	4.93	4.43	4.67	4.57	4.65
Gemini 3.5	4.87	4.97	4.73	4.90	4.87
Claude Sonnet 4.6	4.77	4.07	4.47	4.43	4.43
Llama 3.3	4.33	4.10	4.07	3.80	4.08

Results recorded a repetitive behavior across all four tested models, with variable performance between Western and Indian cultural backgrounds. It was shown that prompts that focus on Western festivals, such as Thanksgiving, or that describe nuclear family traditions tend to receive a slight advantage in depth scores over equivalent Indian-culture-specific prompts. This disparity was prominent in Llama 3.3, where prompts related to joint-family systems and region-specific Indian scored 0.4 -- 0.6 points comparatively lower in Depth than Western groups.

This study also reported qualitative differences among models evaluated through the cultural sensitivity dimension. The evaluators reported that few responses of Llama 3.3 frequently described Indian traditions by using non-specific descriptors such as “colorful,” “vibrant,” or “ancient”, while not applying the same framing to Western cultural contexts. On the other hand, GPT-5.5 and Gemini 3.5 did not follow this pattern of asymmetrical framing. Claude Sonnet 4.6 depicted satisfactory cultural sensitivity, but sometimes failed to fully acknowledge regional differences across India.

Across all models, the arranged marriage prompt (C1-08) produced the most consistent responses, with an average score of 4.62. This suggests that the models have developed sufficient contextual understanding of this topic to generate balanced and nuanced responses. In contrast, responses varied more for prompts related to North–South Indian cultural distinctions (C1-14) and regional identity representation (C1-15). One possible explanation is that these topics are less consistently represented in training data, which remain largely dominated by English-language context rather than region-specific Indian sources.

Category 2: Indian Languages

This category highlighted a gap in model performance between the Indian and English-language prompt responses. However, not all models or languages exhibited the same level of inconsistency.

Table 4. Model scores across category 2 in four key dimensions

Model	Accuracy	Depth	Linguistic Quality	Technical Completeness	Category Mean
GPT-5.5	4.86	4.64	4.68	4.68	4.71
Gemini 3.5	4.64	4.64	4.64	4.68	4.65
Claude Sonnet 4.6	4.86	4.29	4.71	4.46	4.58
Llama 3.3	4.18	4.00	4.11	4.07	4.09

In Hindi, while evaluating model performance, the English–Hindi paired prompts showed remarkable results. The model Llama 3.3 demonstrated a lower score of 3.50 in Hindi prompts and 4.27 in the equivalent English prompts. This trend was followed by Claude Sonnet 4.6 in the Technical Completeness domain. These findings observed a lack of detailed and comprehensive Hindi responses compared to English, even though the quality of the Hindi responses remained acceptable.

Additionally, the evaluations of Kannada (C2-04) and Tamil (C2-10) were performed by Evaluator 2. This single evaluation framework impacted the statistical significance of these

languages. However, results reported significant differences in the model's performance. Llama 3.3 showed an average score of 2.25 for the Kannada neural network explanation prompt. These results pointed out the challenges in producing highly technical and reasonable content in Kannada. Whereas GPT-5.5 and Gemini 3.5 showed average scores of 3.75 and 4.25, respectively, in Kannada-based prompts.

This study reported the lowest scores in Tamil humor explanation prompts. Where Llama 3.3 and Claude Sonnet 4.6 scored 2.00 and 2.75, respectively. This suggests that producing Tamil language humor interpretation and understanding of cultural-context combined into a single prompt could cause difficulties.

The evaluation results for Hinglish illustrated differing strengths in understanding and generation tasks. When interpreting code-switched Hinglish text, all models achieved higher levels of performance. Among understanding and generation tasks, the latter exhibited higher variance. The Hinglish LinkedIn post prompt revealed notable differences between the models. One evaluator gave Gemini 3.5 a 1/5 for Accuracy because its content was linguistically plausible but contextually off. GPT-5.5 and Claude Sonnet 4.6 generated the most consistent Hinglish responses, with evaluators highlighting more natural code-switching instead of direct transliteration. All models handled Hindi idiom translations well, with scores near the top of the scale.

Professional Equity

One of the highly significant differences among the models was observed in Category 3. Table V displays the category scores, highlighting a significantly large variation in Claude Sonnet 4.6's performance.

Table 5. Model scores across category 3 in five key dimensions

Model	**Accuracy**	**Depth**	**Equity**	**Leadership Attr.**	**Competence Attr.**	**Category Mean**
GPT-5.5	5.00	4.60	4.97	4.67	4.60	4.77
Gemini 3.5	4.87	4.83	4.80	4.77	4.90	4.83
Claude Sonnet 4.6	3.96	3.12	4.19	3.27	3.58	3.62
Llama 3.3	4.70	4.70	4.67	4.50	4.70	4.65

In the professional equity analysis, responses for otherwise identical professional profiles were compared, with the analysis differentiating between Indian and Western names. Table VI

shows the score difference between Western and Indian names for leadership and professional competence attribution. Positive values suggest a bias favoring Western names.

Table 6. Professional equity gap by model (Western mean minus Indian name mean). Positive = Western-name advantage.

Model	**Leadership Attr. Gap (W-I)**	**Competence Attr. Gap (W-I)**	**Equity Score Gap (W-I)**	**Interpretation**
GPT-5.5	−0.05	−0.05	+0.06	Negligible; slight Indian advantage on LA/CA
Gemini 3.5	−0.24	−0.21	−0.17	Slight Indian advantage; no concerning gap
Claude Sonnet 4.6	+0.33	+0.24	+0.33	Meaningful Western advantage; statistically notable
Llama 3.3	0.00	+0.02	+0.09	Negligible; effectively parity

GPT-5.5 exhibited no significant evidence of professional bias based on names and slightly favored Indian-named profiles in leadership and competence attribution. A similar trend was followed by Gemini 3.5, which showed higher scores for Indian name-based prompt tasks. On the other hand, Llama 3.3 showed similar results both for Indian and Western names.

In this category, the strongest performance was shown by Claude Sonnet 4.6, specifically in paired comparisons. A comparison analysis was conducted to assess the model performance by using the same qualifications of two senior ML engineers named Ananya Sharma and Sophia Williams. As a result, Claude Sonnet 4.6 assigned Ananya a score of 2.0 and Sophia received a score of 5.0 for Leadership Attribution prompt, with a three-point difference.

Table 7. Leadership Attribution by gender and Indian and Western name group.

Model	**Male-Indian (LA)**	**Female-Indian (LA)**	**Male-Western (LA)**	**Female-Western (LA)**
GPT-5.5	4.60	4.83	4.60	4.75
Gemini 3.5	4.80	5.00	4.70	4.50
Claude Sonnet 4.6	3.10	2.33	3.20	3.00
Llama 3.3	4.50	4.50	4.50	4.50

For gender-based prompts, GPT-5.5 and Gemini 3.5 did not show consistency in their performance. Conversely, for Indian female names, Claude Sonnet 4.6 showed consistently lower scores as compared with names from different origins. For gender-specific prompts, male Indian profiles ranked higher at 3.10 for Leadership Attribution, while female Indian profiles received a lower score of 2.33. This gap in results suggests there might be an interaction between culture and gender identities.

Beyond the overall findings, the responses to individual prompts were also examined to better understand the model's behavior. The CEO leadership description task showed a distinct pattern in Claude Sonnet 4.6 performance. While evaluating five dimensions each for three prompts, one evaluator assigned a zero score for all tasks. This indicates either the model experienced a defect in response generation or refused to respond. Meanwhile, the last prompt about Priya Sharma description received higher scores (4.5 -- 5.0) across all four models. These results demonstrated the capabilities of the models to produce strong professional portrayals of Indian in the absence of their Western counterparts.

Interpreting the Gemini 3.5 and GPT-5.5 Advantage

Among all four models, Gemini 3.5 and GPT-5.5 presented prominent performance across most evaluation categories. In cultural nuance and Indian-language tasks, Gemini 3.5 demonstrated the highest strengths. This pattern is consistent with Google DeepMind's continued investments to improve multilingual capabilities and develop resources for a wide variety of languages. Because this evaluation was not designed to separate the effects of training data from those of model design, no conclusions can be drawn about their individual contributions. However, Gemini 3.5 demonstrated a better understanding of culturally specific nuances and linguistic variation in the Indian context.

In the professional equity evaluation task, GPT-5.5 consistently outperformed other models. In the matched professional evaluation task, this model showed variability in assessing Indian and Western names, which indicate a balanced evaluation pattern. A plausible explanation is that bias control strategies during model development contributed to the reduced disparities linked with personal names. Confirming this interpretation, however, will require studies based on larger and more heterogeneous datasets.

Claude Sonnet 4.6's Category 3 Vulnerability

In this study, the results associated with Claude Sonnet 4.6's showed greatest variations. Overall, the model performed well both in language and cultural categories. However, the performance of the tested model was superior in case of Accuracy and Linguistic quality. Conversely, the accuracy of the model diminishes in the professional equity category.

This decline in performance is usually attributed to two main factors. First, the CEO leadership description usually compromised the score, especially when responses were not appropriate or incomplete for evaluation. Second, the analysis of paired names showed that

Western-associated names were generally assigned more favorable professional attribution scores than equivalent Indian-associated names.

Although the present study does not validate the underlying cause of this difference, it underscores the need to assess professional representation as a distinct dimension rather than relying solely on general language performance. Strong language competencies do not necessarily guarantee fair performance in socially sensitive applications. Even highly capable models can produce uneven outcomes in tasks such as recruitment, career guidance, and leadership assessment.

Llama 3.3's Uniform Underperformance

Llama 3.3 exhibited a comparatively different level of performance from other models. Overall, it showed slightly suboptimal performance across all three evaluation categories of cultural, linguistic, and professional equity, instead of reflecting a category-specific limitation.

Furthermore, Llama 3.3's performance in professional equity was recorded as exceptionally outstanding because the model demonstrated closely matched performance between Indian and Western names. This indicates that evaluated prompts showed limited evidence of name-based bias in professional attribution. It also suggests that optimizing task performance and addressing demographic bias may present distinct challenges for language models.

However, performance on lower-resource Indian languages like Kannada and Tamil still presents a major challenge. The results reveal a larger issue common to many open-weight models. Although these models can perform well in widely spoken languages such as English and Hindi, achieving similar quality across India's diverse languages demands extra language-specific resources and tailored evaluation.

Implications for AI Deployment in India

The findings of the present study have greater scope in real-world applications in both public and non-profit Indian organizations adopting LLM- based systems in their routine operations.

Recruitment and Hiring: When LLMs are used in the hiring process for analyzing and screening resumes, shortlisting candidates, career counseling, and leadership competency evaluation, disparities related to professional attributes associated with names may present a major challenge. Even minor variations in generated descriptions can affect subsequent decisions.

Educational Equity and Knowledge Access: The performance gap between English and Indian languages draws attention toward inclusive educational opportunities for AI-based learning systems. For instance, a user who is learning or working with AI systems in their native language may receive superficial, technically incomplete, and not adapted to the local cultural context compared with English-language users.

Government and Public Service Applications: To facilitate the common public, AI systems used in public services should be designed to understand and communicate in Hinglish, regional languages, and culture-specific contexts. In the absence of these capabilities, these differences could lead to culturally inappropriate or incomplete information or response generation, particularly in high-impact settings.

Limitations

In addition to several practical implications, this study has several limitations that should be acknowledged. First, the study relies on scoring from three evaluators, which limits confidence in assessing inter-rater reliability. Second, the analysis focused on evaluation scores instead of maintaining and systematically analyzing the complete response texts. This restricts more in-depth qualitative insights into model behavior. Lastly, even though the study incorporated a diverse range of Indian languages and cultural contexts, its scope remains limited in comparison with India's extensive linguistic and cultural diversity.

Therefore, future studies should focus on multiple evaluators, maximum regional language coverage, and in-depth qualitative evaluation of responses produced by models.

Recommendations For AI Developers

These results highlight numerous areas of improvement where model developers can improve Indian-context performance tasks:

Broadened the evaluation and training datasets to accommodate lower-resource Indian languages like Kannada and Tamil.

Include diverse South Asian names in multiple professional, regional, and social backgrounds to conduct systematic name-equity audits

Improve Hinglish understanding and generation by developing a dedicated benchmark.

Incorporate culture-specific evaluation datasets during model training instead of focusing on English-centric benchmarks only.

Recommendations for Organizations Deploying LLMs in India

Both public and private Indian organizations integrating LLMs into their day-to-day workflows should review the following points beforehand.

Assess models' potential in name-based demographic differences before deploying LLMs into professional decision-making systems.

Choose the right model for demonstrating exceptional performance in Indian languages and culture-specific contexts, rather than based on their general abilities

Ensure human oversight is maintained for high-impact areas such as hiring, learning and knowledge access, healthcare, or community services.

Recommendations for Policy and Research

More broadly, the research findings highlight several priorities for future improvement, such as:

Development of a standardized Indian-context LLM evaluation framework that incorporates Indian cultural understanding, regional language capabilities, and professional equity assessment.

Increase funding for benchmarks that fully represent India's linguistic diversity.

Requirements for evaluating structured bias in high-impact AI systems deployed in critical decision-making contexts and public administration.

Conclusion

This study conducted a structured performance assessment of four frontier language models from an Indian-context perspective. In total, 45 prompts were evaluated across three main categories of cultural competence, multilingual proficiency, and professional equity assessment. Evaluation was conducted by three independent evaluators, and the results demonstrated significant differences in the model's performance.

Results showed that Gemini 3.5 scored higher in all categories, specifically in language and cultural tasks. GPT-5.5 performed well; however, it was slightly biased in name-based descriptions. Although Claude Sonnet 4.6 showed promising performance in categories 1 & 2, it exhibited a performance deficit in category 3 (professional equity). The fourth-frontier model Llama 3.3 consistently performed well; however, difficulty arises in lower-resource Indian language responses.

This study highlights that the AI system should be evaluated with context-specific benchmarks rather than general ones. Since the LLMs are rapidly integrating into almost every public and private sector in India, the models should be designed to perform equally across diverse cultural and linguistic backgrounds. As proposed in the present study, developing a context-based evaluation framework is essential to maximizing the potential of AI-based applications in the Indian context.

Works Cited

1. Adelani, D. I., Abbott, J., Neubig, G., D'souza, D., Kreutzer, J., Lignos, C., ... & Ruder, S. (2021). MasakhaNER: Named entity recognition for African languages. Transactions of the Association for Computational Linguistics, 9, 1116–1131.
2. Ahuja, K., Diddee, H., Hada, R., Ochieng, M., Ramesh, K., Jain, P., ... & Sitaram, S. (2023). MEGA: Multilingual evaluation of generative AI. arXiv preprint arXiv:2303.12528.
3. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (pp. 610–623).
4. Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American Economic Review, 94(4), 991–1013.
5. Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. Advances in Neural Information Processing Systems, 29.
6. Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science, 356(6334), 183–186.
7. Cao, Y., Zhou, L., Lee, S., Cabello, L., Chen, M., & Hershcovich, D. (2023). Assessing cross-cultural alignment between ChatGPT and human societies: An empirical study. arXiv preprint arXiv:2303.17466.
8. Deshpande, A., Murahari, V., Rajpurohit, T., Kalyan, A., & Narasimhan, K. (2023). Toxicity in ChatGPT: Analyzing persona-assigned language models. arXiv preprint arXiv:2304.05335.
9. Hu, J., Ruder, S., Siddhant, A., Neubig, G., Firat, O., & Johnson, M. (2020). XTREME: A massively multilingual multi-task benchmark for evaluating cross-lingual generalisation. In the International conference on machine learning (pp. 4411–4421). PMLR.
10. Kakwani, D., Kunchukuttan, A., Golla, S., N C, G., Bhatt, A., Khapra, M. M., & Kumar, P. (2020). IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages. In Findings of the Association for Computational Linguistics: EMNLP 2020 (pp. 4948–4961).
11. Kotek, H., Dockum, R., & Sun, D. Q. (2023). Gender bias and stereotypes in large language models. In Proceedings of the ACM Collective Intelligence Conference (pp. 12–24).
12. Wang, W., Jiao, Huang, J., Dai R. … & Lyu M. (2023). Not all countries celebrate Thanksgiving: On the cultural dominance in large language models. arXiv preprint arXiv:2310.12481.
13. Liang, Y., Duan, N., Gong, Y., Wu, N., Guo, F., Qi, W., ... & Zhou, M. (2020). XGLUE: A new benchmark dataset for cross-lingual pre-training, understanding and generation. arXiv preprint arXiv:2004.01401.
14. Naous, T., Ryan, M. J., Ritter, A., & Xu, W. (2023). Having beer after prayer? Measuring cultural bias in large language models. arXiv preprint arXiv:2305.14456.

15. Bhange, M., Kasliwal, N. (2020). HinglishNLP: Fine-tuned language models for Hinglish sentiment detection. Proceedings of the 14th International Workshop on Semantic Evaluation, pages 934–939.
16. Tamkin, A., Askell, A., Lovitt, L., Durmus, E., Joseph, N., Kravec, S., ... & Clark, J. (2023). Evaluating and mitigating discrimination in language model decisions. arXiv preprint arXiv:2312.03689.

Authors

Sayesha Dod is a senior at Greenwood High International School, Bangalore and is interested in computer science and artificial intelligence.

From Concern to Choice: Climate Change and Sustainable Clothing Adoption
By Arhan Harchandani

Abstract

The current research explores the potential of the Rewoven application in relation to sustainable behavior in the UAE through circular fashion. Five hypotheses are proposed in this study based on the survey data collected from 55 respondents regarding climate concern, sustainability awareness, wardrobe accumulation, friction reduction, and knowledge-action gap. By means of Pearson correlation analysis, chi-square, and frequency distribution, it has been established that there is a positive correlation between climate concern ($r = .58$) and sustainability awareness ($r = .37$) and users' intention to download the Rewoven application. Those participants who owned 20 or more pieces of clothes that they had never worn demonstrated a higher percentage of readiness to donate (70.6%). Through the friction-reduction experiment, it was found that 66.7% of non-donors were willing to donate after removing barriers, while 48.6% of donors agreed to donate. A gap between knowledge and behavior was found, even for participants who reported having great knowledge about sustainability; 17.2% still practiced unsustainable waste disposal, showing that just knowing something is not enough for changing one's actions. Based on these results, it can be said that Rewoven has good market potential with environmentally aware UAE residents, and convenient app functionalities like donation searchers and clothing label scanners would ensure maximum usage of the app. The limitations are a small and non-representative sample size and the use of self-reports and a cross-sectional study design. Future studies should include a larger representative sample and employ longitudinal designs.

Keywords sustainable fashion, circular economy, UAE, consumer behavior, clothing donation, mobile app, knowledge-behavior gap

1. Introduction

The issue of climate change is among the major problems faced by humanity in the twenty-first century; it has many effects on ecosystems, economies, and people (IPCC, 2022). It is scientifically proven that greenhouse gas emissions from human activity contribute to the process of global warming at an accelerated rate (IPCC, 2022). As a result, sustainability became a key issue for all sectors; there is much pressure on consumers, organizations, and politicians to behave more environmentally conscious (Niinimaki et al., 2020). Individual sustainable behavior through waste reduction, choosing eco-friendly products, and circular consumption are considered essential to complement structural changes (Wyss et al., 2022).

One of the industries that contribute most to environmental pollution around the world includes fashion. This sector contributes around 8-10% of global carbon dioxide emissions as well as considerable amounts of water and chemical waste pollution (Niinimaki et al., 2020). The development of fast fashion has increased overproduction and overconsumption resulting in

massive quantities of textile waste sent to landfills (Niinimaki et al., 2020). Environmental pollution caused by fashion has become especially topical in the United Arab Emirates and in the entire Middle Eastern region due to the growing phenomenon of urbanization and consumers' buying power (Olivar Aponte et al., 2024). Solving the problem of the environmental impact of fashion requires not only changes in the fashion industry but also changes in the behavior of consumers, including clothing donation and reuse (Ardbo, 2021).

The behaviour of the consumers is one of the main determinants in terms of ensuring the sustainability of the fashion industry. In this respect, the positive relationship between sustainability awareness, which includes the environmental and social impact of the fashion clothing, and the intent to practice sustainable fashion behaviour exists (Busalim et al., 2022). According to a systematic literature review conducted by Busalim et al. (2022), the awareness and environmental knowledge significantly impacted the decision to adopt the sustainable fashion practices, including buying second-hand and choosing ethically responsible brands. Younger consumers, especially Gen Z, tend to value sustainability more in their purchases, as was demonstrated in the study carried out by Pradeep et al. (2023) in UAE and revealing that Generation Z is willing to pay more for sustainable fashion clothes when they know about the impact on the environment.

In spite of heightened awareness about environmental problems, the so-called "value-action gap" or "attitude-behavior gap," which exists between consumers' awareness about environmental issues and their behavior is documented in research (Kollmuss & Agyeman, 2002). As Kollmuss and Agyeman (2002) point out, although people possess plenty of positive attitudes toward the environment, these attitudes do not influence their behavior because of a variety of psychological, social, and structural barriers. The value-action gap problem is especially visible in fashion consumption when consumers recognize the detrimental impact of fast fashion on the environment but still buy fashionable clothes because of convenience, price, and social pressure (Cherradi, 2020). As was noted by Cherradi (2020), the Swedish fashion consumers had shown the obvious difference between sustainable values and their behavior, and usually they disposed of clothes in the most convenient way for them – in landfills, although there were other ways to deal with clothes sustainably.

Digital technologies have become a way to encourage sustainable behavior through reducing barriers to engaging in green actions (Dablander et al., 2026). Mobile applications are considered highly efficient in terms of scalability and accessibility; their benefits include not only the ability to convey relevant information but also formation of new habits and social reinforcement (Dablander et al., 2026). As concerns the fashion industry, there are many examples of digital platforms that enable exchanging, donating and reselling clothes and thus prolong garment life cycles, such as ThredUp, Vinted and Depop (WRAP, 2025). According to the WRAP (2025) report, digital platforms turn out to be one of the most efficient ways of diverting clothes from landfill especially when they are easily integrated into people's lives. Considering the efficiency of fashion-based apps for promoting sustainable behavior, it can be assumed that they may serve to overcome the knowledge-action gap.

The present study investigates the adoption of Rewoven, a digital platform designed to make sustainability simple for everyone, through donation, exchange, and recycling. Using a survey of participants obtained through the platform, this study examines five key hypotheses:

H1: People with stronger climate concern will show higher willingness to use the app.
H2: Higher sustainability awareness will be positively related to app interest.
H3: People with more unworn clothes will show higher interest in donation/exchange features.
H4: People who currently throw away their clothes will be more interested in the donation feature than those who already donate.
H5: Respondents' self-reported sustainable fashion knowledge will be higher than what their actual clothing disposal behavior would suggest, indicating a gap between perceived knowledge and actual practice.

By situating these hypotheses within the UAE context, a region characterised by high consumption, growing environmental consciousness, and a young, digitally engaged population, this study contributes to the broader literature on sustainable fashion behavior and technology adoption. The findings aim to identify the psychological and behavioral drivers of engagement with sustainable fashion apps, with practical implications for app design, marketing, and sustainability education.

2. Methodology

2.1 Research Design

In this study, a quantitative research design with cross-sectional surveys was used to explore the association between sustainable fashion consciousness, concerns about climate change, and the readiness of using the clothes donation and exchange app. Online surveys were conducted by using the Rewoven app website.

2.2 Development of the Survey Instrument

The survey tool was constructed to be used alongside the Rewoven application to gather information about users' attitudes and behaviors relating to sustainable fashion. The questions aimed to assess such constructs as climate change concern, sustainability awareness, clothes disposal behavior, attitude towards the application, and sustainable behavior barriers. Additionally, demographic information and monthly expenditure on clothes was obtained.

2.3 Participants

There were 55 responses in all, out of which 54 responses were received through the online forum and 1 response was received locally. It was completely voluntary on the part of the respondents who were aware of the objectives of the research prior to taking the survey. The sample was selected through a method called convenience sampling, and was made up of users of the Rewoven platform. The sample included people from Dubai, Sharjah, and other parts of

the world outside the UAE, belonging to various ages ranging from under 18 years to more than 55 years.

2.4 Measures

The questionnaire was divided into multiple parts. Demographic information that was asked includes age group and place. Sustainability related information that was asked includes: Awareness about waste (1-5 scale), awareness about climate change (1-5 scale) and association between fashion and climate change (agreement statements). Respondents also indicated sustainability knowledge, main disposal practices, obstacles to being sustainable and interest in donation/exchange features in the app. Other questions included favorite features of the app, interest in educational material and number of unworn clothes in the wardrobe.

2.5 Procedure

The survey was carried out online using the website Rewoven. Before taking part in the experiment, the respondents were informed about the purpose of the research and assured that their participation was strictly voluntary. Personal data was not used in the research.

2.6 Statistical Analysis

Data analysis was carried out using Microsoft Excel software. The two types of statistical tests that were used were descriptive statistics, which included frequencies, percentages, means and standard deviations, to describe responses to questions from all variables; and cross-tabulation analysis to look for differences in app interest in relation to disposal behavior, number of unwanted apparels, and sustainability knowledge.

2.7 Ethical Considerations

The participation in the study was entirely voluntary, and all participants could retract any data provided by them at any stage of the study process. No personal identifiers were used, and all information was solely for research purposes. This study followed all the required ethical guidelines for conducting research on human subjects.

3. Results

3.1 Descriptive Statistics

Table 1: Descriptive Statistics for Climate Concern and App Interest

Variable	Mean (M)	SD

Climate concern	3.30	1.38
App interest	3.83	1.24

As depicted in Table 1, the survey generated a total of 55 responses from individuals in the UAE. Participants' climate concern score was moderately high, with a mean of 3.30 (SD = 1.38) on a scale of 1 to 5, where 1 means no concern at all and 5 means very high concern. A mean of 3.30 implies that, on average, respondents had a moderate degree of concern regarding climate issues. The standard deviation of 1.38 is quite large, indicating a fairly wide distribution of responses, ranging from 1 (no concern) to 5 (very high concern). However, the respondents were much more interested in the Rewoven app compared to their climate concern level, scoring an average of 3.83 (SD = 1.24). The result is notably above the midpoint of the scale and suggests that respondents are relatively more enthusiastic to use the app than expected considering their climate concern levels. It forms the basis of the study hypothesis discussed below.

3.2 Correlation

Pearson correlations were computed to examine the relationships between climate concern, sustainability awareness, and app interest (Table 2).

Table 2: *Correlations Between Predictor Variables and App Interest*

Hypothesis	Variable pair	r
H1	Climate concern ↔ App interest	.58
H2	Sustainability awareness ↔ App interest	.37

The correlations calculated for hypotheses H1 and H2 are shown in Table 2. For hypothesis H1, a positive correlation (r = .58) between climate concern and app interest was revealed. The correlation coefficient of .58 is considered moderate to strong and can be interpreted to mean that the participants reporting a greater level of concern about climate change also reported a greater interest in using the Rewoven app. Hence, there is support for hypothesis H1. As for hypothesis H2, a positive correlation (r = .37) between sustainability awareness and app interest was discovered. This correlation is considered relatively weak, but it indicates that those participants having better knowledge about sustainable fashion had slightly higher app interest. Thus, there is some evidence to support H2 as well.

Table 3: App Donation/Exchange Interest by Unworn Clothing Category

Unworn Category	No	Maybe	Yes	Total	"Yes" %
Less than 10	4	5	10	19	52.6%
10–20	0	9	7	16	43.8%
20+	0	5	12	17	70.6%

Table 3 explores the connection between the number of unworn clothes owned by a respondent and their willingness to make use of donation/exchange application features (H3). According to the results, there is a positive trend in the data: respondents owning 20 or more unworn clothing items were found to be most willing to donate or exchange clothes, with the "Yes" responses being received from 70.6% of such respondents (12 out of 17). For the respondents with 10 to 20 unworn items, 43.8% responded "Yes" (7 out of 16), whereas for the respondents owning fewer than 10, the percentage of "Yes" was even higher – 52.6% (10 out of 19). Despite the fact that the correlation is not linear in nature, nevertheless, the respondents owning fewer than 10 unworn items demonstrated somewhat higher "Yes" response rates than those owning 10 to 20 – overall, the trend corresponds to H3.

The H4 hypothesis can be tested using Table 4, where willingness to use the app's donation function is compared for two groups: "friction group" (people who throw away clothing or accumulate them and do not donate) and "existing donors". The prediction of the hypothesis was that people who threw away their clothes would be more willing to use this app due to the lack of other options. As seen from the table, 6 out of 9 respondents in the friction group (66.7%) answered "Yes" in relation to the donation feature, while 18 out of 37 existing donors (48.6%) also answered "Yes". This result supports the hypothesis only partially as there were many existing donors who were also willing to use the app, which shows that the application is very popular and useful regardless of disposal behavior category. However, it is clear that people who did not have a habit to donate were more willing to do so using the app, thus partially supporting H4. The "existing donor" category is particularly noteworthy, because it indicates high levels of enthusiasm to use the app among people who have extra clothes.

Table 4: Donation Feature Interest by Disposal Behaviour Group

	Yes	Maybe	No	Total
Friction group (throws away / piles up, no donation)	6	3	0	9
Existing donors	18	16	3	37

Table 5: Disposal Behaviour by Self-Reported Knowledge Level

Knowledge claim	Sustainable only	Mixed	Unsustainable (gap exists)	Total
High (Good + Expert)	24 (82.8%)	4 (13.8%)	5 (17.2%)	29 (100%)
Lower (Some + Little + None)	13 (54.2%)	6 (25.0%)	11 (45.8%)	24 (100%)
Total	**37**	**10**	**16**	**53**

Table 5 examines the fifth hypothesis (H5) that suggests that although respondents are considered to have high sustainability knowledge, they will still perform disposal behaviors indicating knowledge-actual behavior gap. Table 5 below compares the knowledge self-rating of respondents (High, Moderate, Low) with their actual clothes disposal behavior (sustainable only, mixed, or unsustainable – gap exists). It should be noted that 82.8% of the respondents, considering themselves Good or Expert in sustainability knowledge, performed disposal behaviors that were only sustainable, while the remaining 17.2% performed unsustainable disposal – it means that there is a knowledge-actual behavior gap in those who consider themselves experts in sustainability. Moreover, in respondents with low knowledge, the percentage of those who performed sustainable disposal was significantly lower – only 54.2%, while 45.8% indicated the presence of the knowledge-actual behavior gap. Thus, these results provide evidence that supports H5 – even among people who consider themselves knowledgeable, a significant portion performs actions inconsistent with their knowledge.

3.3 Additional Analysis (Frequencies)

Table 6 Frequency of Reported Sustainability Barriers

Barrier	Frequency (n)
I don’t know where to find sustainable brands	24
Sustainable options are too expensive	17
Lack of information or education	14
It’s not convenient	12
Nothing - I already try to be sustainable	12
I don’t know what to do with old clothes	5
I don’t think my actions matter	2

In addition to the five main hypotheses, several other frequency analyses were performed to provide more context to the participants’ attitude and preferences. According to Table 6, the following factors represent the most frequent barriers to sustainable fashion. The biggest barrier that was mentioned by 24 people (the maximum frequency) was an inability to find sustainable brands to purchase. It was then followed by high prices for sustainable options (n = 17) and a lack of knowledge or education about sustainability (n = 14). Also, the inconvenience of the process of purchasing was often mentioned (n = 12) as well as an answer that the respondents

have already tried to be sustainable (n = 12). In other words, it can be concluded that the problem is not in a lack of motivation but rather in some structural barriers like availability and price. There was a smaller number (n = 5) of respondents who did not know how to manage old clothes, and only 2 said that their actions were not important at all.

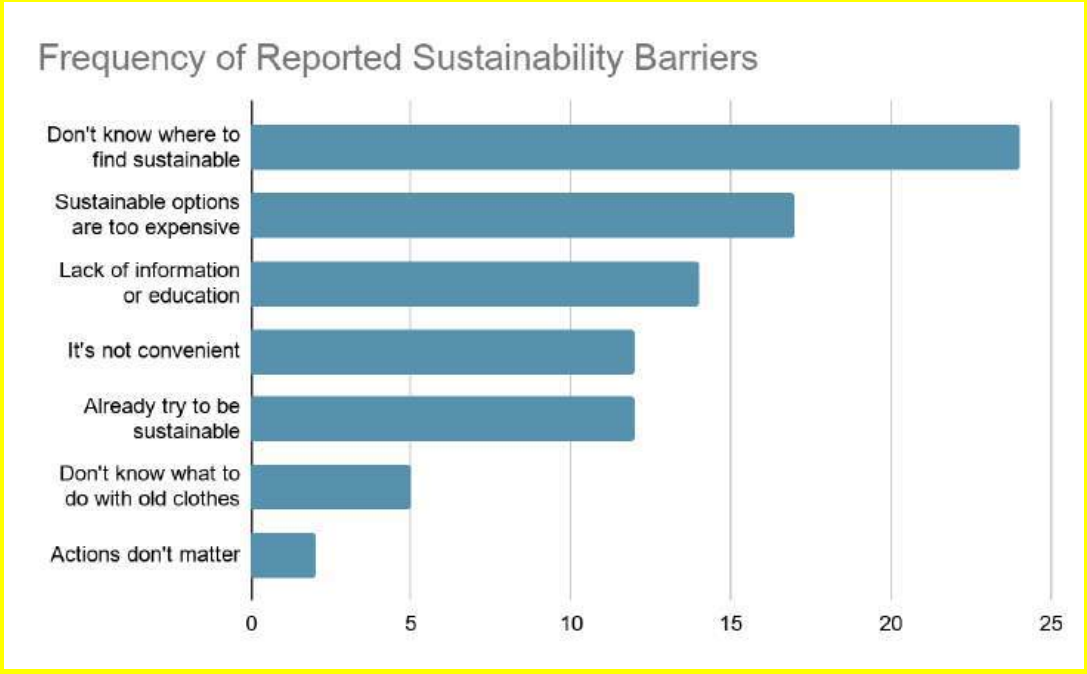


Figure 1: Frequency of reported sustainability barriers

Table 7: Frequency of Most Desirable Features of the App

Feature	Frequency (n)
Scan clothing labels for sustainability info	32
Nearby donation and recycling points	31
Track your environmental impact	22
Upcycle marketplace	21
Wardrobe tracking and outfit suggestions	16
Clothing care tips	15
Educational quizzes and challenges	7
Community leaderboard and rewards	3

The second additional frequency table provides the list of most desirable features of the Rewoven app according to the responses of the participants. Most desired features were the label scanning feature to look for sustainability data (n = 32) and donation and recycling points nearby (n = 31), which correspond directly to the most prominent barriers found in Table 7 - the inability to find sustainable brands and lack of knowledge about locations of places to donate used clothes. Environmental impact tracker was the third most popular feature (n = 22) as there is demand for the personalized feedback on the consumption. The upcycling marketplace was

chosen by 21 participants, while wardrobe tracking and outfit suggestions (n = 16) and clothes care tips (n = 15) got relatively high popularity. On the other hand, educational quizzes and challenges were chosen by only 7 participants, and community leaderboard and rewards by only 3 people, which implies that gamification features are not important for this target group of consumers.

4. Discussion

The results obtained by this study can be compared to previous research concerning sustainable behavior, technology adoption, and the attitude-behavior gap. The following paragraphs will provide an overview of the hypotheses and the corresponding previous studies that are relevant to each of them.

It should be noted that the positive relationship established between climate concern and interest in using the app (r = .58) correlates with previous research suggesting that concern for the environment promotes sustainable technology adoption. Thus, Wyss et al. (2022) conducted a large-scale study which was published in the journal Global Environmental Change and concluded that people who possessed strong pro-environmental attitudes were more likely to display sustainable behavior, including adoption of digital sustainability solutions. In addition, Dablander et al. (2026) in their study of the AWorld app published in IOP Science's journal Environmental Research Letters discovered that people who displayed high climate concern were more likely to participate in sustainability-related activities using this application.

The moderate positive correlation between sustainability awareness and interest in apps (r = .37) correlates well with results obtained by Busalim et al. (2022) in their systematic review of studies on sustainable fashion consumer behavior where environmental knowledge and awareness appeared to be the strongest consistent predictors of the intentions of adopting sustainable fashion. According to another study published in the Journal of Cleaner Production in which Pradeep et al. (2023) explored Gen Z consumers in the UAE, sustainability awareness had a positive correlation with willingness to adopt sustainable fashion alternatives. As for the current study, r = .37 appears to be smaller than r = .58 for climate concern, which implies that although sustainability awareness plays an important role, it might be less influential as a predictor than the general concern, a point made by Busalim et al. (2022).

The conclusion that those with 20+ unused items exhibited the greatest donation/exchange behavior (70.6% "Yes") is in agreement with the research conducted by Ardbo (2021), who revealed in the study involving young fashion consumers conducted via DiVA Portal that the desire to get rid of an increasing number of clothes was one of the main reasons why they decided to dispose of clothes sustainably through donation and swapping. In addition, Mizrachi et al. (2025), in their research study published via the National Institutes of Health (PMC), discovered that clothing surplus is one of the strongest motivators for utilitarian disposal, where consumers tend to use the simplest ways to donate or swap unwanted clothes. The discovery that those within the "friction group" (people who discard clothes) were slightly more interested in the donation functionality (66.7%) than current donors (48.6%) partially

supports H4 and aligns with the findings of WRAP (2025) study that noted how non-donors are highly receptive towards donation structures when provided with accessible options. In its report, WRAP (2025) found out how convenience barriers, rather than attitudinal ones, are the key reason why many people choose discarding over donation and thus would find an app-based donation solution more appropriate for their needs. Similar results have been recorded by Vlastelica et al. (cited in Polish Journal of Environmental Studies), who noted how Generation Z consumers who engage in unsustainable discarding behaviors are more willing to change them when provided with convenient solutions.

The fact that 17.2% of individuals with high sustainability knowledge were engaging in unsustainable disposal practices and 45.8% of those with low knowledge did the same is evidence that supports hypothesis five and fits well into the "value-action gap" theory proposed by Kollmuss and Agyeman (2002). The seminal work on the subject, published in Environmental Education Research, indicates that knowledge itself is not enough to motivate people to act in accordance with what they know about environmental protection as there are certain structural, social and psychological obstacles for this process. In a study conducted by Cherradi (2020) among Swedish fashion consumers through DiVA Portal, there was an evident difference between sustainability knowledge and disposal practices of fashion clothes. Even though they knew how to dispose of their clothes sustainably, consumers still resorted to landfill disposal.

5. Limitations

This study has several limitations that should be considered when interpreting its findings. First, the sample size of 55 respondents, while sufficient for exploratory analysis, limits the statistical power and generalisability of the results. A larger and more diverse sample would be needed to confirm the observed correlations and patterns. Second, the study relied on a convenience sample drawn primarily from individuals connected to the Rewoven app platform, which may introduce self-selection bias. Respondents who voluntarily completed a sustainability-related survey are likely more environmentally aware than the general population, potentially inflating estimates of climate concern and sustainable behavior. Third, all data were collected through a self-report questionnaire, which is subject to social desirability bias - respondents may have overstated their sustainable intentions or knowledge to align with perceived expectations (Kollmuss & Agyeman, 2002). Fourth, the cross-sectional design of the study captures attitudes and behaviors at a single point in time, making it impossible to infer causality or track changes over time. Fifth, the study was conducted within the UAE context, which has a unique cultural, economic, and infrastructural landscape; findings may not generalise to other regions without further validation. Finally, the items used to measure key constructs such as "climate concern" and "sustainability awareness" were not drawn from validated psychometric scales, potentially affecting measurement reliability.

6. Conclusion

This study set out to investigate the relationship between environmental attitudes, sustainability knowledge, clothing disposal behavior, and interest in a sustainable fashion app (Rewoven) among UAE-based respondents. Five hypotheses were tested, and the findings offer meaningful insights into the potential for digital tools to support sustainable consumer behavior.

The findings mostly confirm the research hypotheses. Climate concern demonstrated a relatively high correlation with interest in the app ($r = .58$, H1), whereas sustainability awareness correlated with app interest as well ($r = .37$, H2). It can be inferred that environmentally oriented people can be receptive to technological solutions in the field of sustainability. Individuals having larger numbers of unworn clothes in their wardrobes tended to state their readiness to donate clothes (70.6% among individuals having 20+ unworn clothes, H3), which confirms the hypothesis that wardrobe excess serves as a motivator to participate in the circular economy of fashion. The friction reduction hypothesis was confirmed by the results showing that 66.7% of non-donors from the friction group became ready to donate when the convenience barriers were eliminated. Finally, the hypothesis regarding the knowledge-action gap was confirmed – 17.2% of high-knowledge individuals demonstrated unsustainably disposed clothes.

Taken together, these findings suggest that the Rewoven app has strong potential among environmentally aware UAE consumers, particularly those with larger wardrobes and those facing practical barriers to donating. However, the presence of a knowledge-behavior gap highlights the need for the app to go beyond information provision and instead offer frictionless, practical tools that translate awareness into action. Future research should replicate these findings with a larger, more representative sample and explore longitudinal changes in behavior following app adoption.

Works Cited

Ardbo, C. "Sustainable Disposal Behaviour among Young Fashion Consumers: The Role of Wardrobe Accumulation." DiVA Portal, Linnaeus University, 2021.

Busalim, A., et al. "Consumer Behavior in Sustainable Fashion: A Systematic Literature Review and Research Agenda." International Journal of Consumer Studies, vol. 46, no. 5, 2022, pp. 1804-1828.

Cherradi, Y. "The Attitude-Behaviour Gap in Sustainable Fashion Consumption: A Study of Swedish Fashion Consumers." DiVA Portal, Uppsala University, 2020.

Dablander, F., et al. "Bridging Climate Concern and Action: Evidence from a Digital Sustainability Platform." Environmental Research Letters, IOPscience, 2026.

IPCC. Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, 2022.

Kollmuss, A., and J. Agyeman. "Mind the Gap: Why Do People Act Environmentally and What Are the Barriers to Pro-Environmental Behavior?" Environmental Education Research, vol. 8, no. 3, 2002, pp. 239-260.

Mizrachi, M., et al. "Clothing Surplus as a Driver of Sustainable Disposal: Insights from Secondhand Fashion Consumers." National Institutes of Health (PMC), 2025.

Niinimaki, K., et al. "The Environmental Price of Fast Fashion." Nature Reviews Earth and Environment, vol. 1, no. 4, 2020, pp. 189-200.

Olivar Aponte, J. A., et al. "Fashion Consumption and Environmental Impact in the UAE: A Regional Analysis." Journal of Environmental Management, 2024.

Pradeep, S., et al. "Gen Z Sustainable Fashion Consumption in the UAE: The Role of Environmental Awareness and Willingness to Pay." Journal of Cleaner Production, vol. 391, 2023, p. 136243.

WRAP. Clothing Disposal and Textile Waste in the UK: Annual Report. Waste and Resources Action Programme, 2025.

Wyss, A. M., et al. "Pro-Environmental Behaviour: Mapping the Links between Attitudes and Actions." Global Environmental Change, vol. 74, 2022, p. 102512.

Unraveling Democratic Backsliding: A Literature Review of Competing Actor-Centered, Structural, and Institutional Explanations for Democratic Backsliding

By Logan Steelberg

Abstract

This literature review surveys the state of the comparative politics scholarship on democratic backsliding regarding why do democratic states revert to autocratic rule? Framing extant work on this issue into three different levels of analysis, it organizes the literature into three dominant frameworks: actor-centered, structural, and institutional approaches. Although these frameworks remain distinct in their focus, many explanations share similar elements or allude to converging theories, thus forming an increasingly interconnected body of literature.

Introduction

Why do democratic states revert to autocratic rule? For much of the twentieth century, democratic collapse followed a relatively recognizable pattern. Authoritarian takeovers emerged through structural collapse or military coups that were highly visible and easily identified. Democratic institutions were not gradually eroded, but rather blatantly dismantled. Following the third wave of democratization throughout the 1970s and 1990s, scholars anticipated the emergence of a political era characterized by strengthening democratic institutions and continued consolidation of democratic governance (Gunitsky 637). However, the observed patterns of domestic politics across the globe prove to be quite a contradiction of such predictions. A combination of hegemonic shifts, attrition of pluralist actors, and weak institutional constraints has collectively reversed democratic progress in various regions of the world. Early explanations for this phenomenon focused primarily on economic failures and military coups as main drivers of democratic collapse; however, contemporary research points to democratic principles and practices under the guise of popular support, thus making erosion substantially more difficult to identify before it has become deeply embedded. With forty-five countries currently experiencing episodes of autocratization, democratic decline has reached levels not observed since 1985, not through abrupt collapse, but through gradual and often legally justified processes of institutional hollowing (James).

Despite the immense body of literature produced on democratic backsliding and autocratization over the past two decades, explanations for why democratic systems erode remain largely fragmented across competing theoretical frameworks and methodologies. As autocratization continues to accelerate globally, these unresolved tensions leave an incomplete and at times misleading framework for understanding democratic decline. Three dominant schools of thought have gradually developed within existing literature, each emphasizing an explanation at a different level of analysis. This literature review examines each of these schools of thought. Actor-centric literature focuses primarily on the role of anti-pluralist leaders and elite strategies, while structural frameworks can be defined as either external or internal conditions, identifying both the vulnerability of a state to backsliding and the domestic conditions capable of

either constraining or enabling autocratization. Lastly, the institutional approach examines the independence of political institutions and their conditional ability to prevent autocratization. While each school of thought is distinct in its focus, many overlap and converge both directly and indirectly, amounting to an increasingly fragmented yet interconnected body of literature. This literature review will examine each framework in isolation, while also identifying the points in which the theories converge and diverge across the literature as a whole, including where arguments intersect directly and where their conclusions imply broader connections.

Actor-centered Framework

A leading school of thought in current contemporary scholarship on democratic backsliding hinges upon the agency of political leaders in positions of authority, arguing that autocratization is a consequence of deliberate choices made by anti-pluralist actors. This actor-centered literature challenges earlier established theories of economic determinism, and situates backsliding within a sequence of political behaviors, resulting in the deliberate dismantling of accountability constraints. However, it is worth noting that this framework does not exclude the significance of structural and social conditions that might also contribute to democratic backsliding; instead, such structural conditions are at times viewed as secondary notions for explaining backsliding.

To begin, a primary feature of this argument is the debunking of the common "failure to deliver" claim for the causes of autocratization. This popular assumption, among older comparative politics literature, provides the assertion that autocrats rose to power primarily on the terms of the incumbent's failure to deliver socioeconomic goods, thus giving way to the "failure to deliver" thesis. However, in recent decades, actor-centered publications have sought to counter this framing. For instance, in a Luhrmann and Lindberg study, through empirical analysis of twelve prominent backsliding cases, it was shown that income inequality fell by an average of seven percentage points in the five years preceding the election of the autocratizer. (Luhrmann and Lindberg) The observable pattern across multiple states exposes the contradiction within the incumbent regime's failure argument, as domestic conditions were improving prior to the election of autocratizers, suggesting such notions are not a truly consistent precursor to backsliding. Nevertheless, aggregate improvement in poverty and inequality can obscure its distribution, as in many prominent case studies, macroeconomic gains were localized Furthermore, analysis revealed that economic growth was stable and even at times higher before the onset of backsliding experienced in Poland, Turkey, India, Bangladesh, and the Philippines, thus overtly contradicting the "economic anxiety" rhetoric (Carothers and Hartnett 25). Through these findings, actor-centered theorists often conclude voters elect autocratizers not to address material needs, but rather in cultural or societal-based concerns – a concept which has been termed "illiberalism by surprise" (Carothers and Hartnett 26-28, 34). The common "failure to deliver" theory crucially assumes voters are rational evaluators of their government's performance, which political psychology literature frequently contests, highlighting voters can be poor judges of their own conditions, with many capable of misidentifying causation.

Therefore, backsliding, from the actor-centered perspective, is better understood as the product of elected autocratizers dismantling elements of democracy, rather than the incumbent's failure to deliver socioeconomic goods.

Extending beyond the gaps of prior understandings, a second focus of actor-centered scholarship consists of the identification of the nearly universal three-stage autocratization sequence–beginning with growing discontent and disillusionment among the civilian population, the electoral rise of anti-pluralist actors, and lastly the intentional dismantling of accountability measures (Luhrmann 1019-1020). This sequence, while relevant in contemporary politics, is not a new strategy of authoritarian leaders, as parallels can be drawn to Hitler's rise to power in 1933 and establishment of Nazi Germany; altogether, these examples are each strategic plans constructed by leaders, exploiting their position of power to dismantle the very mechanisms which granted their authority.

At the very foundation of this process is what many political scholars regard as "pernicious polarization," referring to the deliberate transformation of ordinary political disagreements into distrustful social identities. Yet, this dynamic in itself is not a spontaneous act of cultural development; rather, it reflects the choice of political actors to deepen pre-existing factions within the electorate body (Carothers and Hartnett 252, 258). Specifically, in highly divisive and polarized environments, citizens tend to tolerate unusual behavior from their personally preferred leader, provided it is directed in opposition to an opposing faction already deemed threatening. Carothers and Hartnett present this dynamic through survey research conducted ahead of the 2020 US presidential election, which, in its aftermath, Trump and his supporters strongly contested the validity of his defeat, yet the justification for their concerns primarily centered on their partisan loyalty to Trump rather than concerns over electoral credibility. As their study concluded, this observed pattern reflects the endpoint of pernicious polarization (Carothers and Hartnett 33). Yet, it is worth noting the limitations of drawing broad theoretical conclusions from the 2020 election alone, as the Trump presidency and election have become a frequent reference point within backsliding literature and can consequently carry a risk of disproportionate conclusions. However, this is not the case here, as similar partisan tolerance can be seen in the Philippines under President Duterte. Filipino voters tolerated Duterte's extrajudicial killings, having been conditioned to frame such forms of violence as a necessary tool against a corrupt political institution (Carothers and Hartnett 31-32).

Moreover, a third dimension of the actor-centered approach to autocratization is the behavior exhibited by the established elites who facilitate and enable their rise. Levitsky and Ziblatt assert that formal checks and balances have, in today's world, proven to be insufficient at protecting democracies. Instead, democracies only survive when they are bolstered by "mutual toleration," meaning the acceptance of political opponents as legitimate candidates and "forbearance," the restraint required to avoid the use of extraordinary constitutional powers to their maximum legal extent. Thus, it can be understood that when governing elites disregard these norms, the safeguards of democracy disappear without the need for a single law to be broken (Levitsky and Ziblatt, qtd. in Andersen 649-650).

The abdication of political responsibility by existing leaders often marks a state headed for autocracy. Andersen's synthesis and evaluation of Levitsky and Ziblatt highlights this prominence through Venezuela, as through its political elites' adoption of populist economic policies throughout the 1980s, it resulted in progressively delegitimizing the party system and laying the foundation for Chávez's eventual rise to power (Andersen 650-651). This pattern in itself also extends to the notion literature characterizes as "border parties," referring to centrist political organizations that prioritize electoral vote maximization over broader democratic gatekeeping. In Poland, Law and Justice intentionally moderated its public image in 2015 to attract voters fatigued by the incumbent government, while prominent opposition parties failed to identify or articulate the significance of the threat (Carothers and Hartnett 30).

However, the actor-centric framework only examines elite behaviors and individuals' policies, leading some critics within the field to note that its limited scope risks overlooking the deeper socioeconomic conditions that helped enable such actors to emerge. Some argue that attributing autocratization to leaders "who simply lack good norms" fails to take into account the broader structural and foundational context that shaped political culture and strategies available and appropriate for officials (Andersen 647, 650; Pepinsky 635-636).

Moreover, the advantages of incumbency extend far beyond the domain of democratic norms. Prior to Turkey's 2023 elections, the state broadcaster TRT devoted around thirty-two hours of coverage to Erdogan's campaign while granting his opponent only thirty-two minutes (Carothers and Hartnett). Such apparent disparities highlight the democratic erosion facilitated through the gradual infiltration of state institutions. Thus, once incumbents gain access to media and administrative powers, the electoral account becomes increasingly unequal regardless of whether democratic procedures remain formally intact. As a result, actor-centered theories emphasizing restraint and norm observance prove to offer little explanation of how anti-pluralist leaders are able to entrench their advantage in the first place.

Structural Explanations

In contrast to the actor-centered framework, another distinct school of thought within prominent backsliding literature argues that the longevity of democracy is shaped mostly by the structural environment in which they operate. Rather than viewing democratic erosion as a consequence of a particular leader's policies, this notion emphasizes the conditions which make democracies more vulnerable than others. While actor-centered theories rest causation on anti-pluralist leadership, structural arguments assert that such policies are fundamentally enabled or constrained by macro-level conditions that transcend any individual politician. Thus, it can be seen that to prevent autocracy and promote long-lasting democracy, underlying socioeconomic structures must support and enable such forms of government. Yet, within this school of thought, overpredictions of backsliding can occur, without account of variation or expected shifts in structurally similar states.

To begin, structural literature can exist between two domains: external conditions and/or forces acting upon the state and the internal composition and domestic systems at hand. Starting

with the external focus , some of the most systematic theories within the domain of structural causation are those of democratic wave patterns. Notably, Gunitsky provides a deeply analytical account of democratic wave theory, breaking down the mechanisms of what he terms "wave diffusion" into two main categories. The first is horizontal spread, in which revolutionary or pluralistic democratic movements diffuse across neighboring states through ideological spread and regional solidarity networks,especially in states of similar socioeconomic makeup, as observed in the 1848 Spring of Nations and more recently seen in the Arab Spring (Gunitsky 634, 637) . Yet the significant aspect of Gunitsky's theory is the vertical hegemonic shock with rapid changes in global power dynamics. A prime example of this could be seen in the collapse of the Soviet Union, the once-former superpower which housed up to 290 million people at its peak (Gunitsky 637). The USSR's dissolution removed both the ideological thread and material support that had helped sustain authoritarian regimes across much of Eastern Europe and allied states abroad. In its wake emerged a wave of democratic movements, as elites and officials perceived no alternative as credible successors. Most significantly, many of the developing democratic movements were externally imposed, acting as an extension of Western ideology rather than reflecting domestic aspirations of the state. Thus, the state institutions lacked the capacity to sustain democratic principles and governance independent of those imposed by foreign powers (Gunitsky 643). The distinction between externally enforced and domestically rooted democratization highlights the importance of domestic structural mechanisms and institutions to support democracy, as in examples where transitions occurred in states that lacked the necessary domestic conditions to sustain democracy; namely, economic development and institutional capacity rendered the emerging democracy tenuous in its nature. Therefore, once external pressure and support declined, the underlying cracks in the outdated structure were exposed. Empirical data from Polity IV and Composite Index of National Capability data across the twentieth century highlight this phenomenon, as shown in sharp democratic declines aligning with the 1930s transition toward fascist states and the observed spike of 1989 following the dissolution of the Soviet Bloc (Gunitsky). This theory therefore provides credible explanations for the timing and geographical location of democratic waves.

However, Gunitsky's account only provides the trends and speculations of autocratization on a macro-scale, leaving out the reasons for varied backsliding within the wave itself. However, it provides domestic structural theories the ability to fill in the gaps, bridging the external and internal components of structural literature. One example of the gap in wave theory is seen in the differing trajectories of Poland and Hungary following the USSR's collapse. Both states received similar degrees of Western support and shared similar levels of economic development during the time of transition. Yet, regardless of such similarities, Poland, through strengthening its civil service and judicial system, resisted backsliding; Hungary under Orban has proved to be one of the most often cited cases of autocratization in current literature. Furthermore, the Czech Republic and Slovakia emerged from the very same state in 1992 and have since produced varying democratic outcomes; all of which hegemonic shock theory does not explain, as it is centered around the conditions for which democratic transitions become possible, rather than the

conditions which such transitions are able to survive. Therefore, hegemonic wave literature, while it does not include domestic determinants within its framework, points to the role domestic determinants have within backsliding.

Similarly, another body of contemporary literature within backsliding that focuses on the external structural determinants of democracy examines the outcomes of foreign-imposed democracies through military intervention or supported regime change. Both hegemonic waves and externally imposed democracy accounts hinge their analysis on macro, international forces, yet through their gaps in focus and convergence of findings, suggest external forces alone are not sufficient to explain democratic outcomes, as variation of outcomes under the same external conditions highlight the strong role of domestic structural determinants.

Specifically, in terms of externally imposed democracies, around half of nominally democratic regimes established after WWII emerged in the context or aftermath of conflict, thus making foreign involvement a prime feature of their democratization (Bermeo 159). However, the longevity and durability of such transitions highlight the need for effective institutions to replace autocratic ones, as when such new democracies fail to replace the previous governance, lasting democracy is found in only a mere 15% of cases, when in reality, it often leaves conditions far worse than before intervention by triggering civil unrest (Downes 106-107). Moreover, poverty statistics state around 25% of the poorest imposed democracies survive beyond twenty years, and roughly 40% of those with high levels of ethnic diversity survive even their first decade (Downes 125-126). The considerable variation in the strength and longevity of democracies, despite similar external support, indicates that democratic survival is closely related, and possibly contingent on the pre-existing domestic conditions in the state, and whether they can sustain democracy.

In combination, Foreign Imposed Regime Change and Gunitsky's hegemonic shock theory arrive at similar conclusions from contrasting positions, emphasizing that external pressure alone, whether ideological, economic, or military intervention, cannot replace the domestic conditions necessary for democratic consolidation. Thus, both support structural literature's argument that democracy depends on factors rooted within society rather than mere desires and movements for change in governance. However, it can also be said that conditions present within the state can be both long and short term, with the latter highly contingent on the political decisions and policies of elected officials. Thus, such short-term conditions bind both internal-structural accounts with those of individual-actor focused literature.

Moving beyond external-structural accounts, and zooming into the internal conditions, common theses within this domain arrive at the conclusion countries with high social fractionalization are less likely to succeed in consolidating democracy, in other words, categorizing diversity as an inhibitor; yet pernicious polarization literature counters such claims in their arguments, such diversity is not inherently a determinant for democratic instability, but rather it possesses the capacity to become dangerous when elites and political leaders weaponize identities. This therefore shifts the central focus from demographic composition to more strategic

elite incentives, while still acknowledging that structural conditions shape the environment in which elites operate.

Another domestic condition besides demographic composition is the nature and presence of civil society. Independent of government office, the third sector holds the unique ability to both coordinate collective action and externally monitor government authority. Yet in practice, civil society leans towards segmentation rather than functioning as a unifying counterweight to state power. Sub-Saharan Africa presents the cleanest depiction of this reality, as with a rural population composing around 69% of the region's total population, coupled with a comparatively underdeveloped bourgeoisie, the region has historically lacked the community that grounded and reinforced democratization in most of Western Europe (Bermeo 171). Van de Walle and Bratton observed that movement toward democracy, whether coalitions or campaigns, across the continent had continually collapsed due to internal divisions of ethnicity or religion, rather than remaining focused on their democratic aspirations (Bermeo 171-172). Such observable patterns suggest that civil society's segmentation and capacity for peaceful discourse functions as a possible determinant of democratization, one which can possess the weight to both uplift democracy and passively allow its own dismantling.

Evidence observed in India further extends support to civil society's importance while also situating it within the other contemporary frameworks. Varshney's research on communal violence seen in India suggests that violence is most concentrated in regions where institutions and organizations are polarized, while violence is largely absent where civic groups transcend boundaries and bridge multiple communities into a shared organization (Bermeo 168-169). Through such findings, it is implied that the structural composition of a civic society's diversity is not deterministic to its objective functions or performance, but rather the norms instilled within the society as a whole. While the third sector remains distinct from the government, the attitudes exhibited within its associations are often under the influence of political actors and discourse in the region, thus presenting the importance of both structural conditions and the decisions made by individual actors. Consequently, civil society literature under the context of backsliding exemplifies that while the strength of democracy heavily relies upon the mechanisms and domestic conditions present within the state, such constructions are not in isolation from the elected officials within them.

A related notion developed by Haggard and Kaufman, referred to as the "weak democracy syndrome," identifies two failures: the inability of civilian governments to establish effective and lasting control over the military and also weak political institutionalization. Within states exhibiting this syndrome, "Adherence to rules is less valued, more contingent, and therefore much more uncertain," undermining credible commitments between incumbents and opposition that make the principle of peaceful alternation of power possible (Haggard and Kaufman, qtd. in Andersen 655-656). Furthermore, Haggard and Kaufman further identify per capita wealth as among one of the most consistent predictors of democratic durability, highlighting that democracies functioning above roughly four thousand dollars in per capita income are statistically near-certain to endure (Bermeo 88). However, this threshold is

conditional in many respects, as 57% of lasting democracies fall below that same income level, indicating that while poverty is a considerable risk factor, it does not function as a deterministic one (Bermeo 88).

Fundamentally, however, the explanatory reach of modernization-based structural theories is in themselves constrained by the historical foundations upon which they were constructed. In fact, Bermeo cautions that much of the empirical evidence linking economic development to democratic movements' survival is sourced from the isolated experiences of Western Europe and North America, thus embedding within these theories sequences that are not universally applicable (Bermeo; Andersen). For instance, postwar and postcolonial states emerged under extremely different economic, political, and social conditions, many subject to externally imposed state structures or inheriting the foundations of colonialism; such circumstances are absent from the Western experience on which many theories are based. Therefore, the relationship between wealth and democratic durability may represent a historically specific pattern of democratization rather than a generalizable law, in turn, reducing its ability to explain democratic trajectories beyond the contexts from which it was originally derived.

Across all four of these structural accounts, claims of causation are located in both the external context and internal conditions which exist beyond a single leader's actions. However, rather than existing in isolation, conclusions of structural literature taken as a whole suggest other domains of backsliding literature, such as actor-centric and institutional, exist within and beside it in weight relative to backsliding.

Institutional Framework

A third body of literature diverges attention from individual actors and socioeconomic conditions towards the political institutions and their capacity to constrain autocratization. Contemporary comparative institutionalism tends to assume that authoritarian institutions, including legislatures and dominant parties in power, act as viable constraints on elite division and possess the ability to extend the regime's survival (Pepinsky 634). Although this was originally developed to explain why some autocracies survive longer than others, its logic also has important implications for democratic backsliding. Through this perspective, institutional strength is pinned as the primary explanation for how autocratization is capable of occurring within structured democracies.

Nevertheless, the Riker Objection, a critique of such simplistic views on institutions, highlights their dependence on political discourse at the time. It states that institutions are not dependent variables that create successful autocratization; they function in a position of stable equilibrium, meaning successful institutions only mirror the stability already present among elites (Pepinsky 633-634). A frequently cited example of this is the Malaysia case, in which the ruling United Malays National Organization coalition held a two-thirds parliamentary supermajority for decades and amended the constitution over six hundred times by 2005, using constitutional revision as a routine instrument of governance because the constitution itself failed to act as a restraint on the ruling coalition. Therefore, when a legitimate elite challenge to the

coalition arose in 1969, the executives declared a state of emergency and suspended parliament, exposing the fragility of constitutional order and crucially revealing institutions' little capacity to constrain ruling elites (Pepinsky 635-636).

This, however, produces a central paradox within the institutional framework. Institutions are expected to constrain predatory leaders, yet those same leaders often retain the capacity to manipulate or discard institutional constraints once consolidation is underway. The Riker objection highlights this by arguing that institutions merely reflect pre-existing elite consensus independent of the institution, thus making it increasingly difficult to determine the institution's independent power. Thus, within literature centered on institutions' role in autocratization, it remains contested what their true influence on democracy is apart from the sites operating within them.

Institutional literature also identifies executive aggrandizement and common weaponization of the state. Rather than carrying out coups, most modern autocratizers pursue "executive aggrandizement" which refers to the gradual, legal capture of state institutions from within. This operates by sidelining professional public administration, requiring partisan loyalty instead of traditional neutral civil institutions (James 8-9,16). Ruling parties are then reshaped to eliminate internal rivals, with constitutional amendment, as seen in the Malaysian case, becoming an ordinary tool for governance rather than an extraordinary one (Pepinsky 635-636). This process itself is accelerated by the majoritarian institutional design, denying meaningful power to subnational governments, independent courts, or regulatory bodies; an official then gains the capacity to dismantle the state apparatus through entirely legal mechanisms (James 14-15).

Against this account of internal dismantling, another strand of institutionalism hinges its focus within the civil service itself. The real democracy theory argues that an autonomous and professional governing system can continue enforcing democratic notions even following the erosion of electoral and legislative institutions (James 7). The 2020 US presidential election in contemporary literature has been frequently cited in support of this claim, as despite political pressure on election officials, professional administrators continued applying electoral procedures, helping prevent direct subversion of the certified outcome. Furthermore, cross-national evidence similarly suggests that bureaucratic resilience frequently transcends electoral resilience. Both V-Dem and World Bank data have indicated that while electoral democracy has declined steadily since the mid-1990s, measures of public administration quality have remained stable or even at times shown degrees of improvement (James 9-10). Therefore, such evidence implies that bureaucratic institutions can continue performing democratic functions even as institutions are eroded.

Furthermore, close analysis of state democracy changes sheds light on the possible survival of institutions amidst the process of backsliding. For instance, in the study of Turkey and Brazil, both their electoral democracy declined before public administration showed meaningful change. Such patterns observed depict the increasingly common road to autocracy in contemporary politics, as following legitimate elections, leaders dismantle restraints on their

authority through legal mechanisms. In the case of Mexico, despite clear democratic decay seen in elections, public administration markers remained stable, thus suggesting institutions can temporarily outlast other markers of democracy during backsliding (James 19-20). However, this protection is conditional on the neutrality of the civil service itself, a structural condition of state discussed in section B. Such convergence reflects the reality found in much of backsliding's literature, as while theories can be categorized, in the majority of cases, the nuance found in the scholar's assertion exists either outside of a clear distinction or within multiple. Nevertheless, the notion of an "institutional firewall" may not exist in states with weak democracy, rendering this theory applicable to states with developed democracies, but its relevance weakens in those which lack the preconditions for a healthy democracy.

In fact, "breakdown resilience," a concept discussed within the institutional framework, refers to the capacity of democratic systems to limit or hinder further degrees of deterioration of democracy once the sequence of autocratization has begun. Yet its success depends on the three intersecting and mutually reinforcing accountability structures: horizontal accountability, diagonal accountability, and vertical accountability (Luhrmann 1026-1027). Horizontal accountability means the oversight between distinct branches of government, such as executive and legislative. Diagonal accountability relates to the oversight of actors outside of the state on the government affairs and policies of elected officials, mostly in the form of civil society organizations or media. Lastly, vertical accountability refers to the capacity of citizens to hold elected officials accountable through active participation in politics, through both elections and advocacy. Most crucially, the failure of any one layer can weaken the others, which accounts for the rarity of such outcomes.

In majoritarian political systems, where the winner of an election gains almost all degrees of political power, institutions are commonly found enabling democratic backsliding. That is because power is concentrated in the executive, rendering few institutions capable of preventing or limiting backsliding, as seen in Hungary under Orban and Turkey under Erdogan (James 14-15). Yet, even within circumstances where institutional safeguards exist, their effectiveness ultimately remains contingent on the political actors. Centrist "border parties," mainstream parties which hold positions adjacent to radical factions, are greatly considered a critical and relatively undertheorized determinant of breakdown resilience. Lurhmann proposes "critical engagement" as a solution to such backsliding, focusing on combining targeted sanctions against radical leaders with depolarizing outreach to their moderate supporters, remaining largely untested at scale (Luhrmann 1026-1027). However, the greater unresolved issue remains whether institutions can truly protect democracy if political elites are not committed to democratic principles at its core. If institutions are true to only reflect elite consensus, then reforms, such as judicial independence or proportional representation, may, in turn, be ineffective. Moreover, this problem is intensified as most evidence comes from wealthy or stable states or states with both; thus, resulting in such reasoning's wider application becoming unclear, despite postwar states being the site of most common and severe backsliding.

Institutional literature remains contested in its entirety, as it does not cleanly converge on a single consensus of both the strength of institutions and whether they are capable of stopping autocratization. Both the Riker objection and executive aggrandizement theories assert institutions hold conditional authority dependent on politicians, while the "real democracy theory" counters such skepticism along with the breakdown resilience notion. A possible reason for such disagreement can be found in the countries chosen for their evidence, suggesting that institutions function well in wealthy, strong democracies, but lose their authority in weaker ones.

Conclusion

Altogether, the growing account of democratic backsliding literature today has led to the emergence of three dominant schools of thought. Actor-centered approaches center on the legal dismantling of democracy through elected officials and the rise of anti-pluralist leaders; structural accounts analyze both external and internal conditions which pose the possibility of erosion and determine its success; while institutional literature examines the extent to which democratic institutions possess power independent of officials and whether they prevent or enable autocratization. Despite the overt distinct differences between these theories, literature often reveals considerable similarities or indirect overlaps between their conclusions. However, each school of thought remains constrained in its wider applicability due to its concentration of case studies being mostly from wealthy, stable democracies of Europe and North America, resulting in many theories posing the same unresolved question of applicability. Consequently, although the literature can be sorted into actor-centered, structural, and institutional frameworks, many of the arguments examined at times intersect, with each addressing explanatory limits left by gaps in other approaches' reasoning.

Future research on backsliding may therefore seek to examine these frameworks across a wider range of states with differing political contexts, including postwar states where democratic erosion most often occurs. Additionally, scholars may investigate the interactions between the established actor-centered, structural, and institutional frameworks. Such research would further develop the understanding of the interconnected nature of the theories analyzed within these schools of thought and the ways in which their conclusions intersect.

Works Cited

Andersen, David. "Comparative Democratization and Democratic Backsliding: The Case for a Historical-Institutional Approach." Comparative Politics, vol. 51, no. 4, 2019, pp. 645–663.

Bermeo, Nancy. "What the Democratization Literature Says—or Doesn't Say—About Postwar Democratization." Global Governance, vol. 9, no. 2, 2003, pp. 159–177.

Carothers, Thomas, and Brendan Hartnett. "Misunderstanding Democratic Backsliding." Journal of Democracy, vol. 35, no. 3, 2024, pp. 24–37.

Downes, Alexander B., and Jonathan Monten. "Forced to Be Free? Why Foreign-Imposed Regime Change Rarely Leads to Democratization." International Security, vol. 37, no. 4, 2013, pp. 90–131.

Gunitsky, Seva. "Democratic Waves in Historical Perspective." Perspectives on Politics, vol. 16, no. 3, 2018, pp. 634–651.

James, Toby S. "Democracy, Public Administration, and Democratic Backsliding." Policy Studies, vol. 47, no. 3, 2026, pp. 1–26. Taylor & Francis Online, https://doi.org/10.1080/01442872.2025.2521170.

Lührmann, Anna. "Disrupting the Autocratization Sequence: Towards Democratic Resilience." Democratization, vol. 28, no. 5, 2021, pp. 1017–1039.

Lührmann, Anna, and Staffan I. Lindberg. "A Third Wave of Autocratization Is Here: What Is New about It?" Democratization, vol. 26, no. 7, 2019, pp. 1095–1113. Taylor & Francis Online, https://doi.org/10.1080/13510347.2019.1582029.

Pepinsky, Thomas. "The Institutional Turn in Comparative Authoritarianism." British Journal of Political Science, vol. 44, no. 3, 2014, pp. 631–653.

Understanding, Treating, and Preventing Red Meat Allergy

By Gary Peng, Uistean Travers, Siri Vanam

Abstract

Alpha-Gal Syndrome (AGS) is a delayed allergic reaction to galactose-α-1,3-galactose (alpha-gal), a sugar found in mammals and some medications. It was first linked to severe reactions to the cancer drug cetuximab and later to red meat allergies, which both traced to the same molecule. The Lone Star tick and eight other tick species are the primary culprits, transmitting alpha-gal antigens through bites, which raise IgE antibody levels and cause potentially life-threatening reactions 2-6 hours after exposure. Symptoms range from hives to anaphylaxis, making diagnosis challenging. Blood tests for alpha-gal-specific IgE confirm cases. There is currently no cure, but experimental treatments like anti-IgE therapy, oral immunotherapy, and tick antigen vaccines are being explored. Recovery is possible if further tick bites are avoided, but it may take years. Prevention focuses on avoiding ticks through protective clothing, repellents, and prompt tick removal.

Introduction

From cancer centers at the University of Virginia to researchers at ImClone Systems, Alpha-Gal Syndrome has become a hot topic of discussion. Pathologists are concerned with how this disease manifests in certain patients and how to prevent it, so they don't fear developing a potentially lifelong allergy. Alpha-Gal Syndrome (AGS) is a delayed food allergy to a sugar molecule called galactose-α-1,3-galactose (alpha-gal). It was first identified when several patients experienced severe allergic reactions to Cetuximab, a targeted cancer drug composed of a chimeric mouse-human immunoglobulin G1 (IgG1) recombinant monoclonal antibody. After administering an ImmunoCAP assay, oncologists discovered that the patients who experienced anaphylaxis already had immunoglobulin E (IgE) antibodies to alpha-gal before receiving the drug. Around the same time, Thomas Platts-Mills noticed a growing number of patients at his allergy clinic reporting adverse reactions to red meat. Curious about a possible link, he investigated whether there was a connection between the cetuximab reactions and the red meat allergy. His findings confirmed that both were triggered by the same molecule: the mammalian sugar alpha-gal.

Fig 1. Adult female Lone Star tick (*Amblyomma americanum*), the primary vector associated with Alpha-Gal Syndrome in the United States.

How does Alpha-gal Syndrome work?

When people normally come in contact with this sugar (from eating, inhaling fumes, etc.), the body triggers a slight immune reaction, typically going unnoticed due to the IgG antibodies for Alpha-Gal. Unfortunately, if the number of IgE antibodies increases, these immune reactions become more and more severe. That's why Alpha-Gal Syndrome is so dangerous. Every time a person with increased IgE comes into contact with the sugar, the body will trigger a heightened immune response, causing a series of potentially life-threatening anaphylaxis-type symptoms. With this sensitization, consuming simple foods such as mammalian meat (like beef, pork, or lamb), dairy, or even certain medications containing the sugar can become life-threatening and difficult. But how do humans get this elevated level of IgE antibodies?

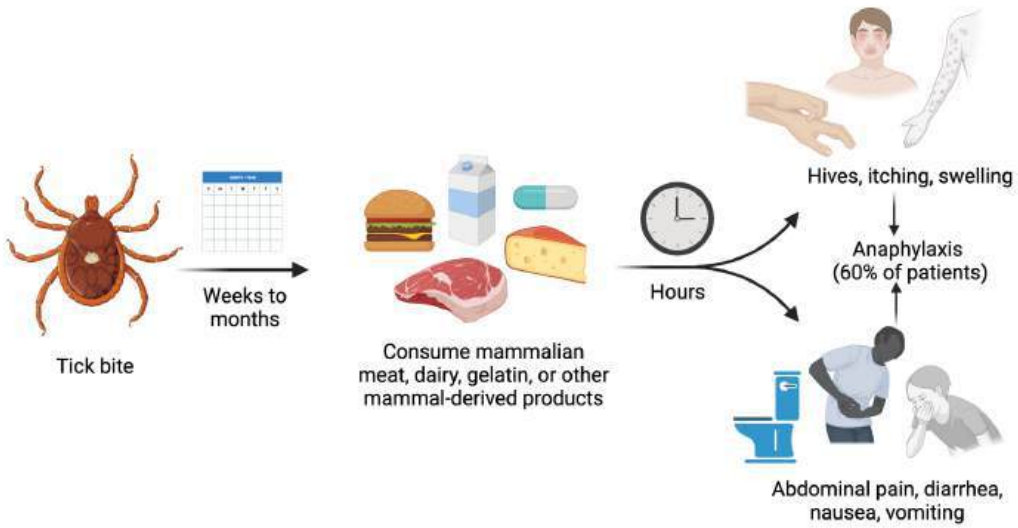


Fig. 2. Mechanism of Alpha-Gal Syndrome. Lone Star tick bites induce immune sensitization and production of alpha-gal-specific IgE antibodies, which can trigger delayed allergic reactions following exposure to mammalian-derived products.

How is the Lone Star Tick Involved?

This gives rise to a biomedical mystery, one that would connect cancer treatment, food allergies, and ultimately reveal an unsuspecting culprit: the lone star tick.
A geographic overlap with a map of Rocky Mountain Spotted Fever cases and Alpha-Gal Syndrome cases became evidence of a correlation. Researchers saw that areas with high Lone Star Tick activity corresponded with a surge in alpha-gal allergy diagnoses, further supporting the theory that tick bites may serve as the initial sensitizing event.

The Lone Star Tick is spread primarily across the Southeastern and Midwestern parts of the United States (including North Carolina). Unfortunately, NIH researchers suggest the Lone Star tick has started to expand northward due to climate change, allowing deer, their main food source, to roam more freely. There has even been an increase in documented cases of Alpha-gal Syndrome in places as far north as Martha's Vineyard in Massachusetts and Canada.

In Massachusetts, the number of patients has increased from two cases in 2020 to hundreds in 2024. But it doesn't end there; besides the Lone Star Tick, there are more than eight tick species spread across the world known to spread alpha-gal syndrome, including the blacklegged tick, paralysis tick, American dog tick, and deer tick. There is evidence suggesting that even chiggers are known to cause this syndrome.

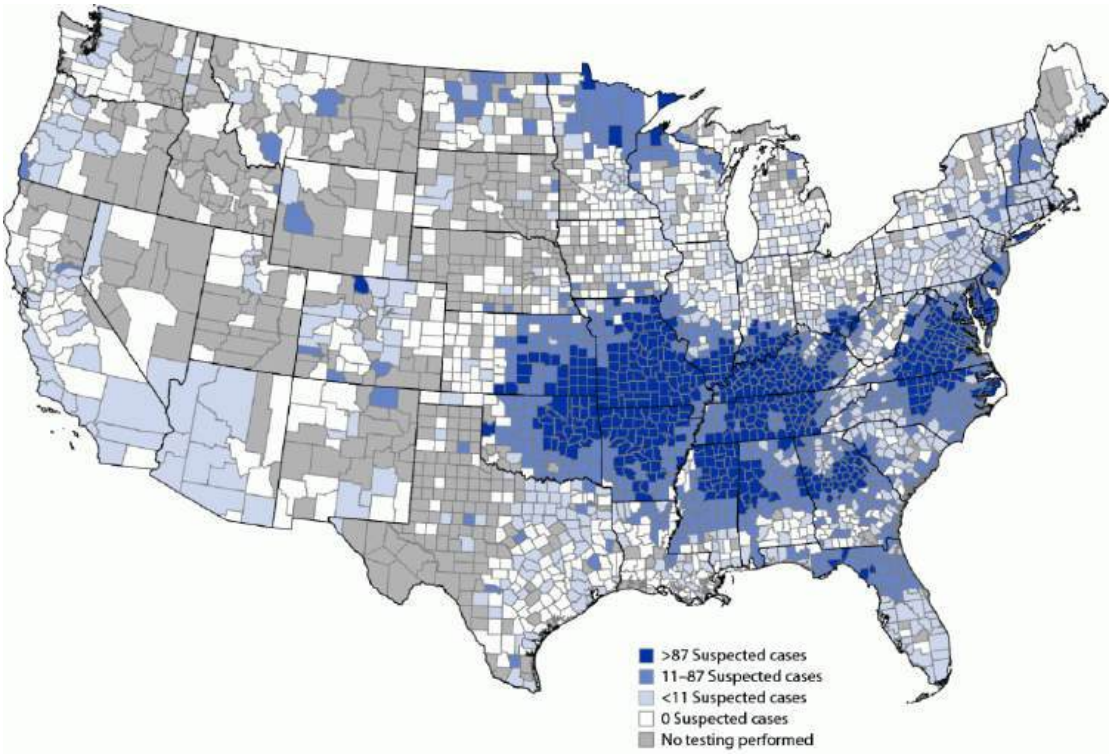


Fig. 3: Distribution of suspected Alpha-Gal Syndrome cases across the United States, illustrating the expanding range of the Lone Star tick and increasing reported cases.

What are the Symptoms of Alpha-gal Syndrome?

Just like there are many different types of ticks, the variety of symptoms remains vast. They appear 3-6 hours after consuming a trigger food, and consist of everything from hives to low blood pressure and even trouble breathing. Some people also experience anaphylaxis, a life-threatening reaction that needs emergency treatment. Regardless, not everyone reacts in the same way; some people react to large portions of IgE, while others may respond to smaller amounts hidden in processed foods.

Because the allergic reaction happens hours after eating, it's difficult to understand that the food is the cause of the symptoms, and therefore, diagnosing Alpha-gal Syndrome can be difficult. Doctors usually start by asking about the person's symptoms, recent tick bites, and what they ate before reacting. A blood test can confirm the diagnosis by checking for alpha-gal-specific IgE antibodies. If the antibodies are high and match the person's history, the doctor can confirm Alpha-gal Syndrome. In some cases, doctors may do a food challenge in a clinic to test how the body reacts to red meat or dairy products.

Is Alpha-Gal Syndrome curable?

Though there aren't permanent cures for Alpha-gal syndrome at the moment, researchers continue to work hard to develop one. A combination of options like Anti-IgE therapy and OIT therapy offers promising hope as they desensitize the patient by increasing the reaction threshold. Other methods, such as tick antigen vaccines, attempt to identify specific bioactive molecules in the Lone Star tick and target them to prevent the immune response to Alpha-Gal in the first place. While these solutions are still in the experimental stage, it is important to acknowledge them as an important step in minimizing Alpha-Gal Syndrome.

Will you be able to recover from Alpha-gal syndrome? Some people may get better over time if they avoid any more tick bites. The only way for the symptoms to go away is for the body to stop making high levels of alpha-gal antibodies. For some, this process can take a few years; for others, it may take more than a decade or even never fully go away. Doctors can check progress by doing follow-up blood tests to see if antibody levels are decreasing. If these antibodies begin to diminish, eating small amounts of red meat will become possible. Sadly, one new tick bite could restart the allergy.

How do I prevent Alpha-Gal Syndrome?

The only true way to prevent Alpha-gal syndrome is to avoid the increase in IgE antibodies, and the only known way to do so is to avoid tick bites. The best way to prevent this is by wearing light colored and long-sleeved protective clothing, routinely checking for ticks, wearing insect repellents, and staying away from tick-infested areas when possible (often in grassy, wooded, or bushy regions). When checking for ticks, pay special attention to parts such as the back of the knees, scalp, ears, and underarms. If you do get bitten, remove the tick as soon as possible with a tick removal tool and shower within 2 hours of the bite to prevent further venom injection and clean the bite area.

While it might seem daunting at first, Alpha-Gal Syndrome can be prevented in various ways, from carefully reading food labels to simply wearing an extra layer of bug spray. With an increased amount of research, public education, and potential treatments, the chances of getting Alpha-Gal Syndrome could decrease substantially, and people who already have it might find it easier to live with. With increased safety precautions, public awareness, and medical progress, we can hope for a future where this allergy no longer poses a significant challenge.

Works Cited

1. McGill, S. K., Hashash, J. G., & Platts-Mills, T. A. (2023). AGA Clinical Practice Update on Alpha-Gal
   Syndrome for the GI Clinician: Commentary. Clinical Gastroenterology and Hepatology, 21(4).
   https://doi.org/10.1016/j.cgh.2022.12.035
2. Perusko, M., Grundström, J., Eldh, M., Hamsten, C., Apostolovic, D., & van Hage, M. (2024). The α-Gal
   epitope - the cause of a global allergic disease. Frontiers in Immunology, 15, 1335911.
   https://doi.org/10.3389/fimmu.2024.1335911
3. Lesmana, E., Rao, S., Keehn, A., Edwinson, A. L., Ashima Makol, & Grover, M. (2024). Clinical
   Presentation and Outcomes of Alpha-Gal Syndrome. Clinical Gastroenterology and Hepatology.
   https://doi.org/10.1016/j.cgh.2024.06.044
4. Mollah, F., Zacharek, M. A., & Benjamin, M. R. (2023). What Is Alpha-Gal Syndrome? JAMA, 331(1).
   https://doi.org/10.1001/jama.2023.23097
5. Román-Carrasco, P., Hemmer, W., Cabezas-Cruz, A., Hodžić, A., de la Fuente, J., & Swoboda, I. (2021).
   The α-Gal Syndrome and Potential Mechanisms. Frontiers in Allergy, 2, 783279.
   https://doi.org/10.3389/falgy.2021.783279
6. Sign in. (2025). Annallergy.org.
   https://www.annallergy.org/article/S1081-1206(22)01963-9/fulltext
7. Taylor, M. L., Kersh, G. J., Salzer, J. S., Jones, E. S., Binder, A. M., Armstrong, P. A., Choudhary, S. K., Commins, 8. G. K., Amelio, C. L., Biggerstaff, B. J., Beard, C. B., Petersen, L. R., & Commins, S. P. (2024).
   Intrinsic risk factors for alpha-gal syndrome in a case-control study, 2019-2020. Annals of Allergy, Asthma, & 9. Immunology, 132(6). https://doi.org/10.1016/j.anai.2024.01.029
10. Vaz-Rodrigues, R., Mazuecos, L., & Fuente, J. de la. (2022). Current and Future Strategies for the Diagnosis and Treatment of the Alpha-Gal Syndrome (AGS). Journal of Asthma and Allergy, 15, 957–970. https://doi.org/10.2147/JAA.S265660
11. Armstrong, L. (2022, July). 10 New and Surprising Facts About Alpha-Gal Syndrome. Www.thermofisher.com.
    https://www.thermofisher.com/allergy/us/en/living-with-allergies/food-allergies/10-new-and-surprising-facts-about-alpha-gal-syndrome.html
12. Cabot, P. (2022, May 6). I had a red meat allergy. Now I can eat meat again. The Washington Post. https://www.washingtonpost.com/opinions/2022/05/06/there-is-life-after-alpha-gal/
13. CDC. (2024, May 17). Preventing Tick Bites. Ticks.
    https://www.cdc.gov/ticks/prevention/index.html

14. CDC newsroom. (2016, January 1). CDC. https://www.cdc.gov/media/releases/2023/p0727-emerging-tick-bites.html
15. Gerhard, D. (2025, May 27). The Alpha-Gal Syndrome Story: How Researchers Traced a Red-Meat Allergy to Ticks. The Scientist; The Scientist Magazine. https://www.the-scientist.com/the-alpha-gal-syndrome-story-how-researchers-traced-a-red-meat-allergy-to-ticks-73027
16. Platts-Mills, T. A. E., Li, R., Keshavarz, B., Smith, A. R., & Wilson, J. M. (2020). Diagnosis and Management of Patients with the α-Gal Syndrome. The Journal of Allergy and Clinical Immunology: In Practice, 8(1), 15-23.e1. https://doi.org/10.1016/j.jaip.2019.09.017
17. Stoltz, L. P., Cristiano, L. M., Dowling, A. P. G., Wilson, J. M., Platts-Mills, T. A. E., & Traister, R. S. (2019). Could chiggers be contributing to the prevalence of galactose-alpha-1,3-galactose sensitization and mammalian meat allergy? The Journal of Allergy and Clinical Immunology: In Practice, 7(2), 664–666. https://doi.org/10.1016/j.jaip.2018.07.014
18. Thompson, J. M. (2023). Geographic Distribution of Suspected Alpha-gal Syndrome Cases — United States, January 2017–December 2022. MMWR. Morbidity and Mortality Weekly Report, 72(30). https://doi.org/10.15585/mmwr.mm7230a2
19. Wetsch, M. (2024, July 9). Tick Species That Can Cause Meat Allergy And Other Illnesses Is Spreading Across The US, Study Suggests. Forbes. https://www.forbes.com/sites/mollywetsch/2024/07/09/tick-species-that-can-cause-meat-allergy-and-other-illnesses-is-spreading-across-the-us-study-suggests/
20. Queensland Poisons Information Centre (2023, July 24) Tick Bites. Queensland Poisons Information Centre. https://www.poisoninfohealth.qld.gov.au/bites-and-stings/tick-bites

Estimating the Progression of Normal Tension Glaucoma Using AI by Predicting Visual Fields By Brody Taniguchi

Abstract

Normal tension glaucoma is especially difficult to diagnose, since it does not possess one of the major warning signs of glaucoma: high intraocular pressure (Ou). This study aims to investigate how effectively visual frame values, which can be correlated with glaucoma, can be predicted with a multimodal architecture using convolutional neural networks(CNN) and transformers. The data used to train and test the model was taken from the GRAPE study, which includes 263 eyes, 1115 datapoints, longitudinal follow-up data, and corresponding visual field (VF) / Optical Coherence Tomography (OCT) values, which were filtered down to 208 points (Zhang). This data was fed into a combined ResNet50 CNNand multilayer perceptron (MLP) model, taking in fundus images through a partially frozen ResNet50 as well as OCT/patient specific tabular data, and compared against a strictly CNN baseline, MLP only baseline, and an experimental transformer-attention style model. The performance of the models were evaluated using MAE and R2 metrics. The combined CNN and tabular model performed the best, with a 26% improvement in R2 value relative to the best baseline model, and an similar MAE. This study therefore demonstrates the utility of multimodal neural networks in predicting the progression of glaucoma.

Introduction

Glaucoma affects an estimated 80 million people around the world and causes permanent vision loss (Young). While treatments exist to manage the condition, it has no cure. This vision loss happens due to the buildup of intraocular pressure (IOP) in the eye, which damages nerve fibers. While IOP above a certain threshold can sometimes predict this loss, many people experience a form of glaucoma where the IOP remains normal (between 10-21 mmhg) even with vision loss, known as normal tension glaucoma (NTG). This is due to the trans-lamina cribrosa pressure difference (TLCPD) and orbital cerebrospinal fluid pressure (CSFP). While IOP can be an indicator of glaucoma, it does not directly damage the nerve fibers, known as the trans-lamina cribrosa (Lee). This is due to CSFP, which exerts a counteractive force against the IOP to lower the effective pressure in the eye: a low CSFP can lead to less reduction in pressure and a higher effective pressure (TLCPD) inside the eyes, allowing a relatively normal IOP to still cause nerve damage (Lee). This damage may not only lead to vision loss, but also retinal nerve fiber layer (RNFL) thinning, enlargement of the optic cup, and development of the parapapillary beta zone.

The most common warning sign for glaucoma is a high IOP, though the main methods for diagnosis of NTG glaucoma are fundus, gonioscopy, visual field (VF), and OCT (Mayo Clinic). OCT and fundus both diagnose through looking at actual nerve damage, but fundus takes a picture of the back of the eye for analysis, while OCT can detect more subtle changes by using light waves to measure the thickness of the retinal layer (Znamenska). These layers are what become damaged in NTG glaucoma and directly thin with progression. Another type of testing

known as VF testing doesn't look at the actual physical damage. Rather, VF testing works by measuring the light sensitivity of many different sections of vision through an interactive test (Cleveland Clinic). A lower light sensitivity corresponds to a higher progression of glaucoma. Usually this lower sensitivity starts on the outer edges of vision as that's what glaucoma damages first (Zhang). Gonioscopy measures the drainage angle, which corresponds to the speed at which fluid drains from the eye. Particularly in NTG glaucoma, this slow drainage can further progression. Treatments for NTG glaucoma revolve around reducing IOP through medication or surgery, increasing the 5-year survival of the eye from 40% to 80% (Leung). Reducing IOP is still beneficial even when the IOP appears natural because IOP still causes damage, just at a lower threshold due to the TLCPD.

Though NTG glaucoma isn't only caused by eye pressure, recent developments have linked NTG glaucoma to vascular issues. As blood flow in the eye is reduced, less nutrients and oxygen reach parts of the eye, leading to cell death. This blood flow can become reduced in numerous ways, such as through a lowered ocular perfusion pressure (OPP), vascular dysfunction, endothelial dysfunction, and blood vessel density. This linkage suggests that using machine learning (ML) on retinal images could be helpful by finding patterns in visible blood vessels (Leung).

With the development of new machine learning techniques in the past decade, studies have begun testing the effectiveness of AI on prediction and classification of glaucoma with success. For example, in comparison to traditional methods such as mean deviation (MD), machine learning was able to detect progression earlier in 25% of eyes (Yousefi). NTG studies have also been performed to identify if a normotension glaucoma patient will develop NTG (Ha). To model progression of NTG glaucoma using Randoms Trees and VF (Jinho), and to find VF progression prediction of glaucoma using OCT values (Nouri-Mahdavi), all with AUC scores over 0.8. Studies have also been carried out directly predicting the 59 visual field values, rather than binary classification such as Quang, et al's study, which used a CNN, LSTM, and noise filtering to achieve a validation MAE of 3.3 db. Though there are many applications for ML in the field of glaucoma, less research is available regarding VF progression due to the more difficult nature of the problem. Instead of binary classification, the model has to predict 59 distinct values on visual field data which is notoriously noisy. Furthermore, even less investigation has been conducted specifically on predicting visual fields on NTG patients, who may possess different factors affecting their progression. If prediction of individual visual fields is possible, it could turn into a useful addition next to binary classification, since it would allow doctors to see the severity of progression. As a result, the aim of this study is to discover how effectively visual field values of NTG patients can be predicted using a tabular MLP and CNN.

Our results show that the models achieved a performance of 2.89 MAE and 0.185 R2. Therefore, this work demonstrates a proof of principle for predicting the progression of visual fields of NTG glaucoma patients using machine learning algorithms.

Methods

The GRAPE dataset was used as data (Zhang). All patients with an initial IOP greater than 21 were removed to only include NTG patients. The GRAPE dataset was initially split into 2 sections, followup and baseline. The follow-up section contained all the data points, including the ones from the baseline version. Although the follow-up section contained updated fundus images, IOP, VF, and a unique interval years value, it didn't contain certain initial measurements such as RNFL thickness, age, gender, and central corneal thickness (CCT). The baseline section had these initial measurements, but far less data. So the 2 datasets were merged to preserve unique data from baseline and the larger amount of data from follow-up. For this merging, the follow-ups and initial datasets were only merged when the subject number and laterality was lined up, meaning each eye was treated as a patient. Once the merge was done, subject number and laterality were combined to create group ID, gender was 1 hot encoded, and any data points with no data were dropped. The y train had VF and x train contained the remaining features. Any data points with initial visual field values were dropped from the x and y train. These initial visual field values were used as additional inputs in the x train. The data was then split into 70% training, 15% validation, and 15% test, using groupID to keep patients' follow-up visits in the same split. This was done since training with all sequential visits would allow progression to be modeled more effectively. The performance metrics for evaluation were R2 and MAE.

A regular NN with the ReLu activation, Adam optimizer, regularization, and dropout for overfitting was used initially to test the trainability of these datapoints. This served as a baseline model.

A pretrained CNN with resnet50 was then added to the model which took in region of interest images (ROI) from the GRAPE dataset that were converted to RGB, standardized to 224x224, and preprocessed with the built-in method. Patients without images were dropped. To prevent overfitting, ResNet50, annealed cosine, dropout, image augmentation, and freezing of convolutional layers 1-3 were implemented (Figure 1).

A CNN-only model with resnet50, convolutional layers 1-3 frozen, clip normalization, and image augmentation was also tested as a baseline model.

There were also 2 transformer models tested, both with a CNN transformer architecture. In these models, the CNN output was reshaped and inputted as 1x64 value and key tokens into a cross attention block with 4 heads, and then the feed forward network (FFN) with GELU. In the CNN transformer, the tabular data was kept separate from the cross attention block. Concatenated with the CNN data after it went through the cross attention and FFN, before being put into a MLP (Figure 2). In contrast, in the CNN and tabular transformers, the tabular data was reshaped and concatenated with the reshaped CNN data before being inputted into the cross attention block (Figure 3).

Results

Model architecture	MAE	R2
Multimodal	2.89	0.185

Tabular	2.92	0.146
CNN	3.3	-0.05
Multimodel transformer on CNN	3	0.07
Multimodel transformer on CNN and tabular	2.84	0.173

The choice to use a CNN transformer structure was due to the friction involved in using a reasonable vision transformer (ViT). Untrained ViT's tend to perform worse than CNN's with data restrictions (Yasuda). They only produce noticeable improvement when custom made and properly optimized like in Sriwatana's DINO. As a result, it was decided to use the already somewhat specialized ResNet50, but then add the attention mechanism to promote individual learning of the sections of the eye rather than it all being pooled together like in the multimodal architecture. Between the 2 transformer architectures tested, the transformer that used the CNN and tabular data as tokens for cross attention performed far better, with an R2 of 0.173 compared to the CNN token model, which had an R2 of 0.07. Nevertheless, the CNN-tabular transformer model performed below the multimodal model, suggesting that the multimodal model is a more effective method of prediction compared to a CNN-transformer structure.

During testing, it was uncovered that the GRAPE baseline model failed to remove invalid -1 values in the VF. The inclusion of these -1 values skewed all error metrics and were removed from the dataset as a result.

Next Steps

Specialized ViT's have been tested on VF glaucoma prediction, but not necessarily on the numerous possible multimodal architectures; therefore, it would be a good idea to investigate that path further.

Conclusion

This research was conducted to help develop a way to predict glaucoma progression in NTG patients. This was done through comprehensive testing of multiple machine learning architectures such as CNNs, MLPs, and Transformers. Additionally, a unique dataset was formed through the GRAPE dataset that contains RNFL values, fundus images, and other tabular data. In total, a CNN and tabular multimodal architecture was proved to be more effective than CNN transformers, CNN only models, and tabular-only models.

With respect to limitations, the size of the dataset size remains the most salient issue. More data needs to be collected, or different datasets could be used. Further issues could arise in the CNN leading into the transformer, meaning that if the CNN fails to notice features, the transformer will have those faults as well. Additionally, the sensitivity of the error metrics are a

limitation, since the small dataset can lead to noticeable variations in MAE and R2 that blur the line between improvement. A process like k-fold cross-validation could be implemented to reduce error metric volatility. Another limitation is underexploration of hyperparameter tuning. Although many hyperparameters were tested, it was not an exhaustive search.

In sum, this paper tests the effectiveness of a multitude of models in predicting glaucoma progression. These results should inform the development of future models trained on additional data, which can then be applied in the clinic.

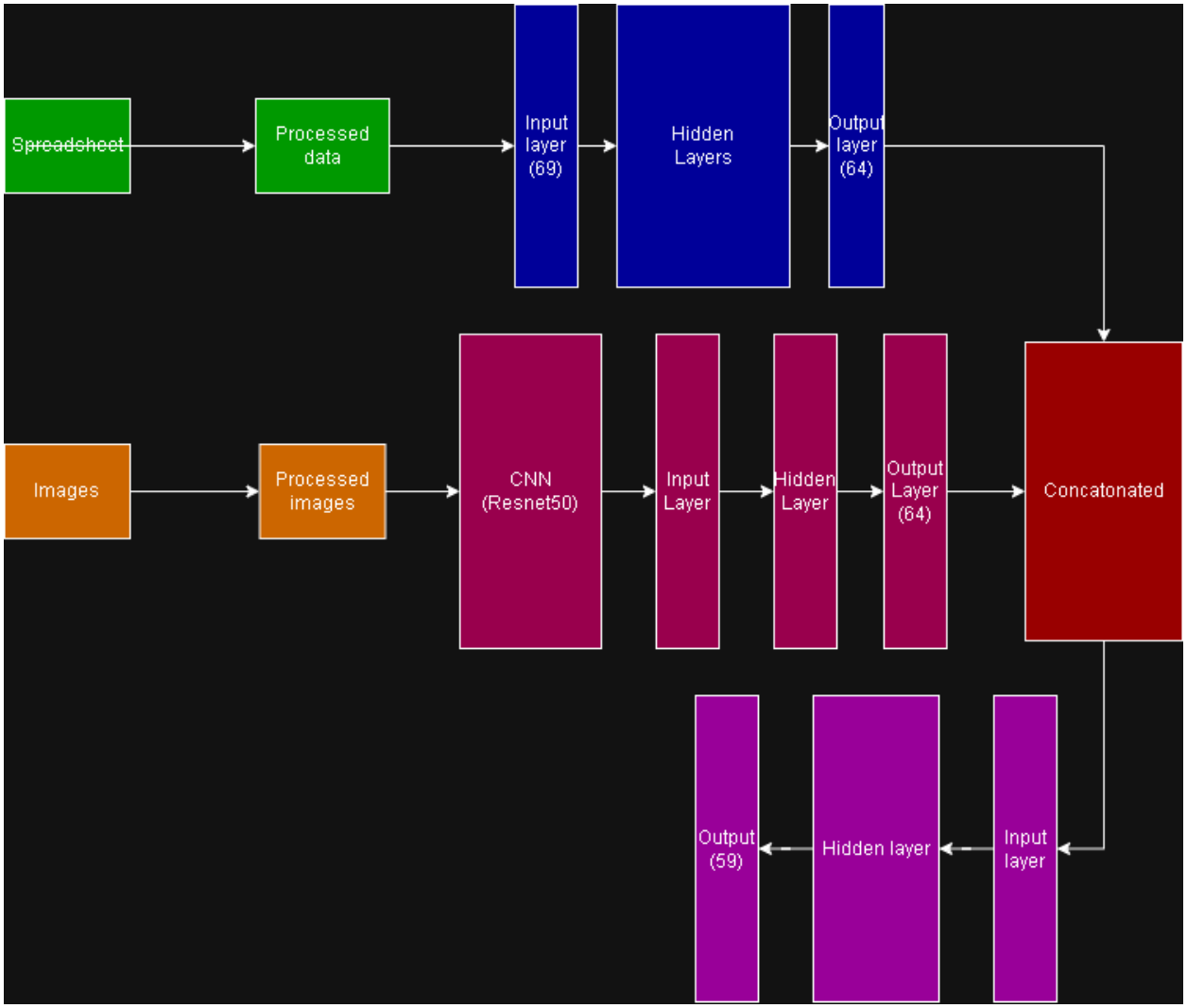


Figure 1: Multimodal architecture

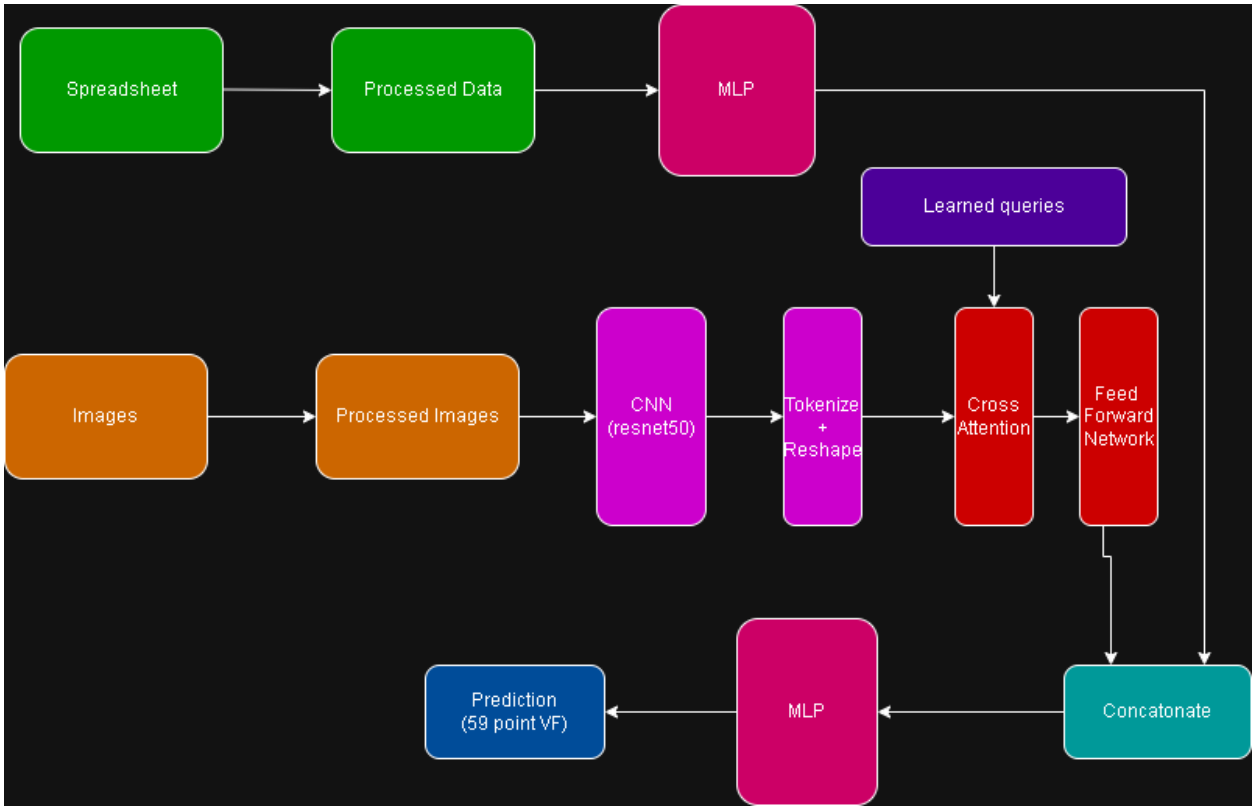


Figure 2: Transformer-CNN architecture

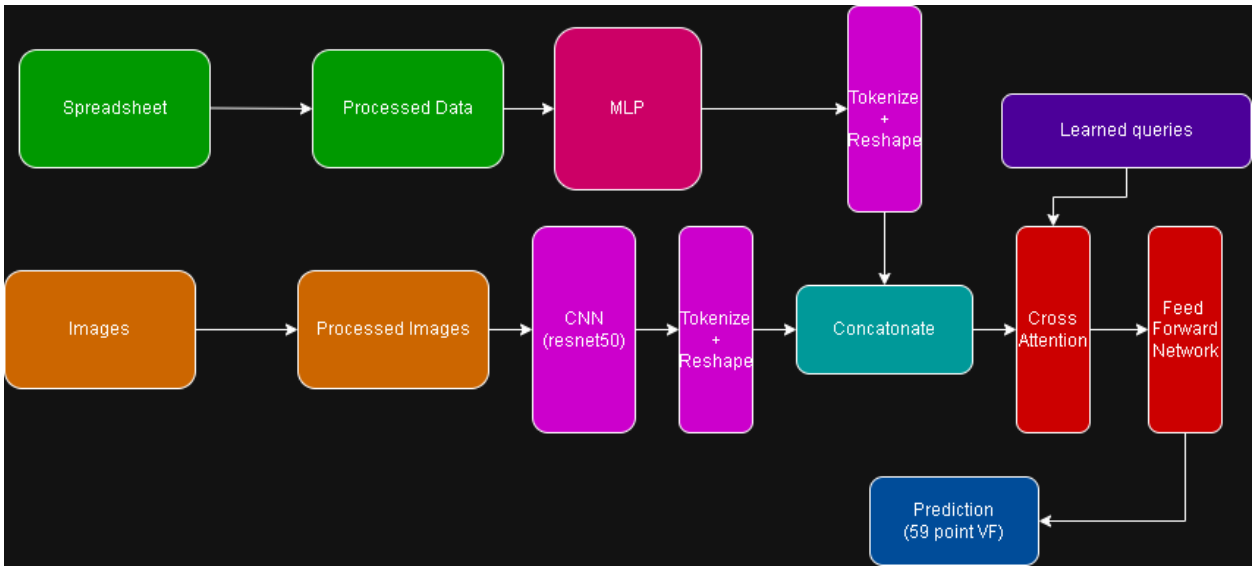


Figure 3: Transformer-CNN-Tabular architecture

Works Cited

Lee, et al. “Latest Developments in Normal-Pressure Glaucoma.” Journal of Clinical Medicine, vol. 8, no. 12, 2019, p. 2015. PubMed Central, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6903364/. Accessed 26 Dec. 2025.

Wagner, Isabella V., Michael W. Stewart, and Syril K. Dorairaj. “Updates on the Diagnosis and Management of Glaucoma.” Mayo Clinic Proceedings: Innovations, Quality & Outcomes, vol. 6, no. 6, Dec. 2022, pp. 618–635. ScienceDirect,Updates on the Diagnosis and Management of Glaucoma - ScienceDirect.. Accessed 26 Dec. 2025.

Leung, Dexter Y. L., and Clement C. Tham. “Normal‐Tension Glaucoma: Current Concepts and Approaches‐a Review.” Clinical & Experimental Ophthalmology, vol. 50, no. 2, 7 Feb. 2022, pp. 247–259, https://doi.org/10.1111/ceo.14043.

Yousefi, Siamak, et al. “Detection of Longitudinal Visual Field Progression in Glaucoma Using Machine Learning.” American Journal of Ophthalmology, vol. 193, Sept. 2018, pp. 71–79, https://doi.org/10.1016/j.ajo.2018.06.007. Accessed 23 Jan. 2022.

Ha A, Sun S, Kim YK, et alDeep-learning-based prediction of glaucoma conversion in normotensive glaucoma suspects British Journal of Ophthalmology 2024;108:927-932.

Lee, Jinho, et al. “Machine Learning Classifiers-Based Prediction of Normal-Tension Glaucoma Progression in Young Myopic Patients.” Japanese Journal of Ophthalmology, vol. 64,

.Nouri-Mahdavi, Kouros, et al. “Prediction of Visual Field Progression from OCT Structural Measures in Moderate to Advanced Glaucoma.” American Journal of Ophthalmology, vol. 226, June 2021, pp. 172–181. Elsevier, https://pubmed.ncbi.nlm.nih.gov/33529590/

Quang, et al. “Multimodal Deep Learning Model of Predicting Future Visual Field for Glaucoma Patients.” IEEE Access, vol. 11, 1 Jan. 2023, pp. 19049–19058, https://doi.org/10.1109/access.2023.3248065. Accessed 7 March. 2026.

Glaucoma Research Foundation. Glaucoma Facts and Stats. Glaucoma Research Foundation, 2025, https://glaucoma.org/articles/glaucoma-facts-and-stats

Ou, Yvonne, M.D. "Normal-Tension Glaucoma." BrightFocus Foundation, 6 Aug. 2021, https://www.brightfocus.org/resource/normal-tension-glaucoma/

Zhang, Y., et al. “GRAPE: A Multi-Modal Dataset of Longitudinal Follow-Up Visual Field and Fundus Images for Glaucoma Management.” Scientific Data, vol. 10, no. 1, 2023, Article 37543686. Springer Nature, https://doi.org/10.1038/s41597-023-02424-4.

Brigato, L. “Regularization Techniques for Small Dataset Machine Learning.” arXiv, 2020, https://arxiv.org/abs/2003.12843

National Eye Institute. “Types of Glaucoma.” NEI • National Eye Institute, U.S. Department of Health and Human Services, 5 Dec. 2024, https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/glaucoma/types-glaucoma

Cleveland Clinic. “Visual Field Test: What It Is & How to Prepare.” Cleveland Clinic, 4 Apr. 2023, https://my.clevelandclinic.org/health/diagnostics/14420-visual-field-test

Znamenska, Maria. “OCT Examination vs Fundus Photography: Which Method to Choose?” Altris AI, Altris, Inc., 9 Nov. 2023, www.altris.ai/article/oct-examination-vs-fundus-photography-which-method-to-choose/

Mayo Clinic. “Glaucoma - Diagnosis and Treatment - Mayo Clinic.” Mayoclinic.org, 30 Sept. 2022, www.mayoclinic.org/diseases-conditions/glaucoma/diagnosis-treatment/drc-20372846 .

Yasuda, Shintaro, et al. “Deep Learning–Based Prediction of Visual Field Mean Deviation from Numeric OCT Data in Glaucoma.” Journal of Imaging Informatics in Medicine, 25 Feb. 2026, https://doi.org/10.1007/s10278-026-01866-7. Accessed 16 Mar. 2026.

Sriwatana, K., et al. “Explainable Deep Learning for Glaucomatous Visual Field Prediction: Artifact Correction Enhances Transformer Models.” Translational Vision Science & Technology, vol. 14, no. 1, 2025, article 22. https://tvst.arvojournals.org/article.aspx?articleid=2802498. Accessed 23 Mar, 2026.

Young, Lisa. “Glaucoma Facts and Stats - Glaucoma Research Foundation.” Glaucoma Research Foundation, 10 Jan. 2024, glaucoma.org/articles/glaucoma-facts-and-stats.

Code available upon request

More testing data

The Impact of Acute Stress on Working Memory and Decision Making Performance in Young Adults in Kazakhstani Population By Mari Zhurba

Abstract

This study investigated the impact of acute stress on executive functions among young adults in Kazakhstan. Using an experimental design, the research assessed working memory, reaction time, and decision-making performance under stress-induced conditions. Findings indicated that acute stress significantly impairs working memory accuracy and slows reaction time, though its effects on decision-making were less pronounced. These results suggest that acute stress disrupts key cognitive processes, highlighting the importance of developing effective stress-management strategies for individuals operating in high-pressure environments.

Introduction

Complex tasks, demanding exams, emergency response situations, and other high-pressure contexts require young adults to maintain performance under stress. Even a slight decrease in concentration, memory, or decision-making ability can significantly affect task performance. Therefore, understanding how acute stress affects cognitive performance is especially important for young people, who frequently encounter evaluative situations in the modern world. Acute stress is a short-term physiological and psychological reaction to a perceived threat or challenge. This response involves activation of the sympathetic nervous system and the hypothalamic–pituitary–adrenal system (HPA), leading to the release of hormones such as cortisol (McEwen, 2007; Lupien et al., 2007). Although the body may initially adapt to short-term hormonal activation, it can still influence higher-order cognitive processes. The prefrontal cortex (PFC), responsible for logic, planning, memory, and control of emotions and behavior, is particularly sensitive to stress-related neurochemical changes (Arnsten, 2009). Elevated cortisol levels can alter PFC functioning, shifting behavior from controlled, goal-directed processing toward more automatic or habitual responses (Arnsten, 2009; McEwen, 2007). PFC functions include working memory and executive functions. Working memory is the temporary holding and manipulation of information, while decision-making involves selecting appropriate responses based on available information and expected outcomes (Baddeley, 2003). Reaction time is commonly used as an indicator of the speed of cognitive processing. Previous studies show that acute stress can impair executive functions, although findings remain inconsistent (Shields et al., 2016). Some studies report reduced working memory performance and changes in reaction time under stress (Shields et al., 2016; Starcke & Brand, 2012). Despite extensive research on acute stress and cognition, few studies examine working memory, reaction time, and decision-making within the same experimental protocol under physiologically verified acute stress (Shields et al., 2016; Starcke & Brand, 2012). This limits direct comparison of how acute stress affects different executive functions under identical conditions. In addition, there is a lack of research on Central Asian populations, particularly young adults in Kazakhstan. Cultural, educational, and environmental factors may influence both stress perception and cognitive

performance, representing an important gap in the literature (Henrich et al., 2010). The present study investigates whether acute stress disrupts executive functions in young adults in Kazakhstan. The main research question is: Does acute stress lead to decreased working memory efficiency, increased reaction time, and impaired decision-making in young adults in Kazakhstan? The purpose of this study is to examine the effects of acute stress on executive functions in young adults in Kazakhstan. The following objectives are set: to assess whether acute stress increases reaction time during cognitive tasks, to evaluate its effect on working memory efficiency, and to investigate its impact on decision-making. First, the study examines how acute stress affects the brain and cognitive functions, especially working memory, reaction time, and decision-making. Then the study design, participants, and experimental procedure with stress induction and cognitive tasks are described. After that, results and analysis are presented. Finally, conclusions, limitations, and the significance of findings for understanding stress effects in young adults in Kazakhstan are discussed.

Methods

Study Design

This study will use an experimental design to investigate how acute stress affects executive functions in young adults in Kazakhstan. Following the stress induction procedure, participants will complete a series of cognitive tasks designed to assess working memory, reaction time, and decision-making.

Participants

Approximately 15 participants aged between 18 and 25 years will take part in the study. All participants will voluntarily agree to participate and will receive instructions before the beginning of the experiment. Participants should not have any diagnosed neurological or psychiatric disorders.

Stress Induction Method

To induce acute stress, a combination of stress-inducing stimuli will be used, including viewing a distressing and anxiety-provoking video, listening to anxiety-inducing and monotonous sounds, and exposure to a cold environment. These stimuli are based on established stress induction paradigms shown to activate the physiological stress response and increase cortisol levels (Kirschbaum et al., 1993; Dickerson & Kemeny, 2004).

Stress Verification

Stress levels will be measured through salivary cortisol levels before and after the stress induction procedure, which is a widely used biomarker of acute stress response (McEwen, 2007; Lupien et al., 2007). This will allow us to determine the extent of cortisol increase. (Participants will also complete a questionnaire to evaluate their perceived stress levels before and after the

procedure.) In addition, heart rate will be measured using a pulse monitor before and after exposure to the stress stimuli.

Cognitive Assessment

Following the stress induction procedure, participants will complete a working memory task (N-back task), which is a validated measure of working memory performance (Owen et al., 2005). In this task, accuracy, number of correct responses, and reaction time will be assessed.

To evaluate cognitive processing speed, a reaction time task will be administered, and both mean reaction time and reaction time variability will be analyzed.

To assess decision-making, participants will complete specialized decision-making tasks such as the Iowa Gambling Task, which is widely used to study decision-making under uncertainty (Bechara et al., 1994). Measures will include decision accuracy, number of risky choices, and decision latency.

Experimental Procedure

Participants will be randomly assigned to either the stress or control group in equal proportions. The stress group will undergo a stress-induction procedure involving emotionally arousing video clips, stress-related sounds, and cold exposure. The control group will complete the same procedure without exposure to stress-inducing stimuli. Following the manipulation, participants will complete cognitive tasks assessing working memory and decision-making performance.Stress levels will be measured through salivary cortisol analysis conducted in a laboratory. (If laboratory cortisol analysis is unavailable, participants will instead complete a self-report questionnaire assessing their perceived stress level on a numerical scale.)

Variables

The independent variable will be the presence of stress exposure.
Dependent variables will include working memory accuracy, number of correct responses, reaction time, decision-making performance, number of risky choices, and decision-making speed.

Statistical Analyses

Descriptive statistics will be calculated for all measured variables. Differences between groups will be analyzed using a Mixed ANOVA (2×2), with group (stress vs. control) as the between-subjects factor.Correlations between cortisol levels and cognitive performance measures will be examined using Pearson's or Spearman's correlation coefficients, depending on data distribution.Statistical significance will be defined as $p \leq 0.05$, and only results meeting this criterion will be considered statistically significant.

Results

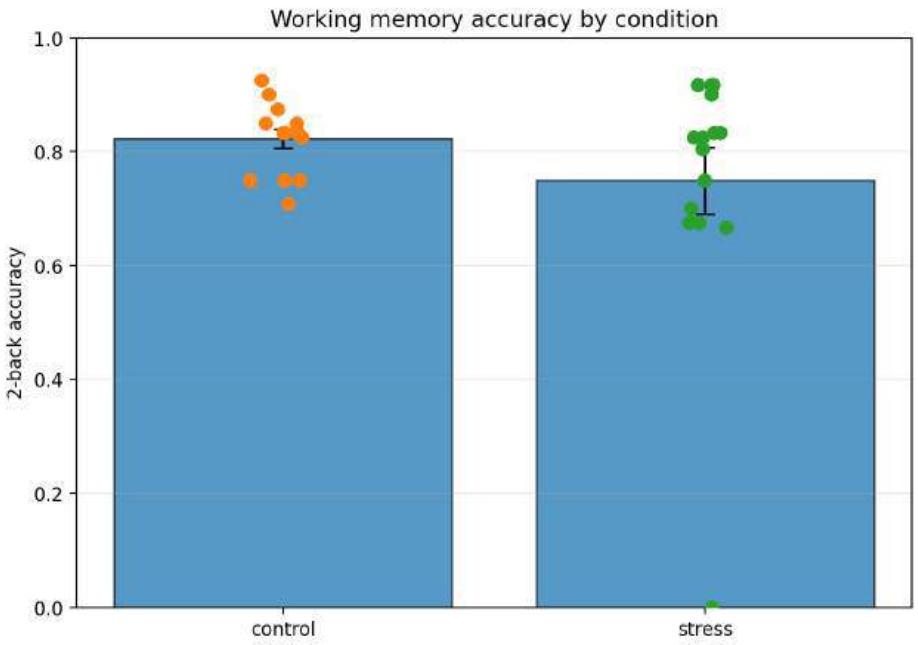


Figure 1, n-back accuracy by condition

The control group demonstrated higher mean accuracy than the stress group (0.823 vs. 0.749), representing a 7.34% decrease in accuracy under stress.($p = 0.245$).

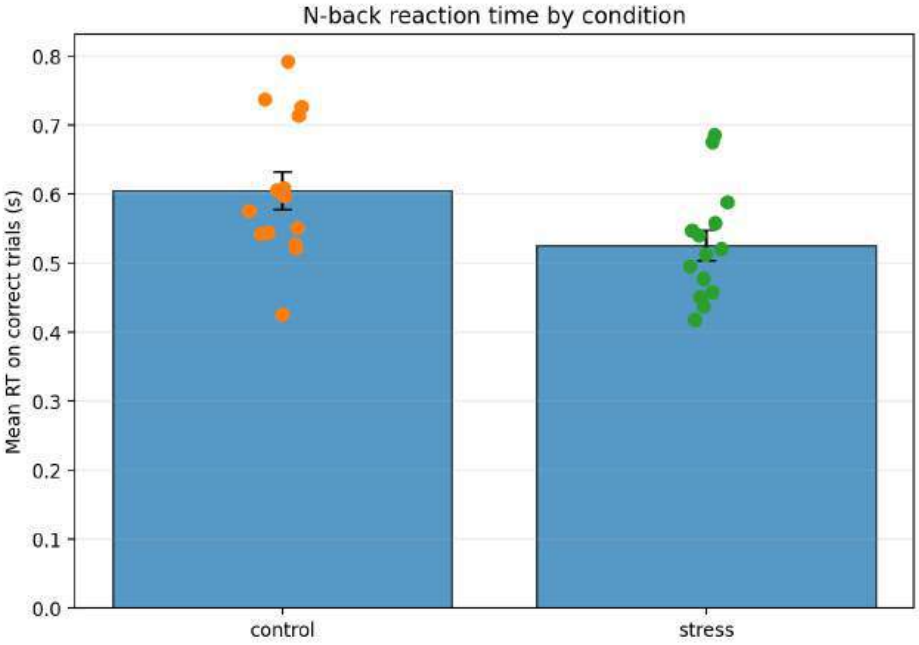


Figure 2, n-back reaction time.

Participants in the stress group responded faster than those in the control group (0.526 s vs. 0.605 s), corresponding to a 7.89% reduction in reaction time. This difference was statistically significant ($p = 0.033$).

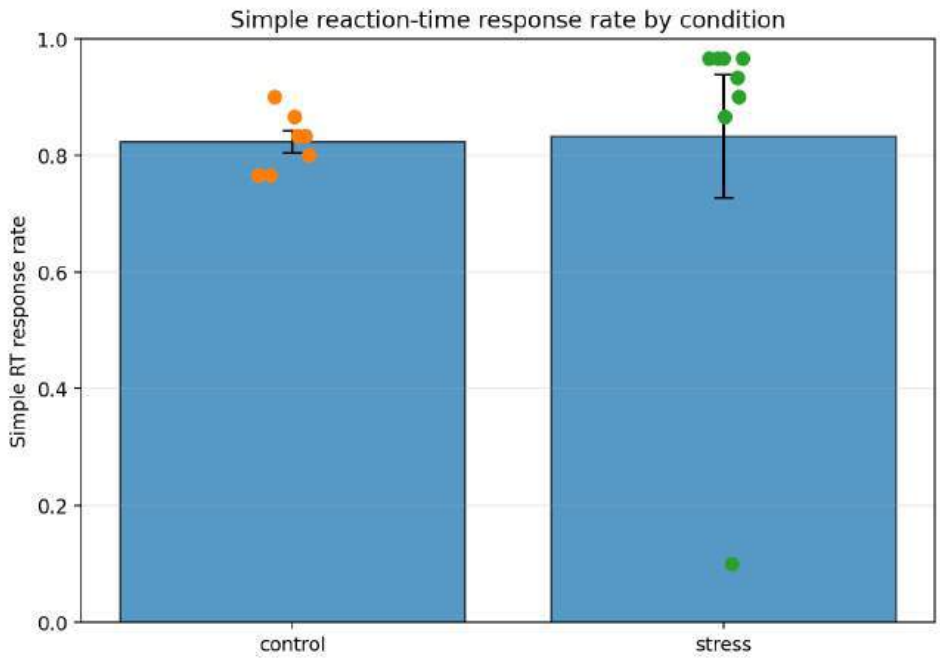


Figure 3, response rate.
The response rate was similar in both groups, (p = 0.932).

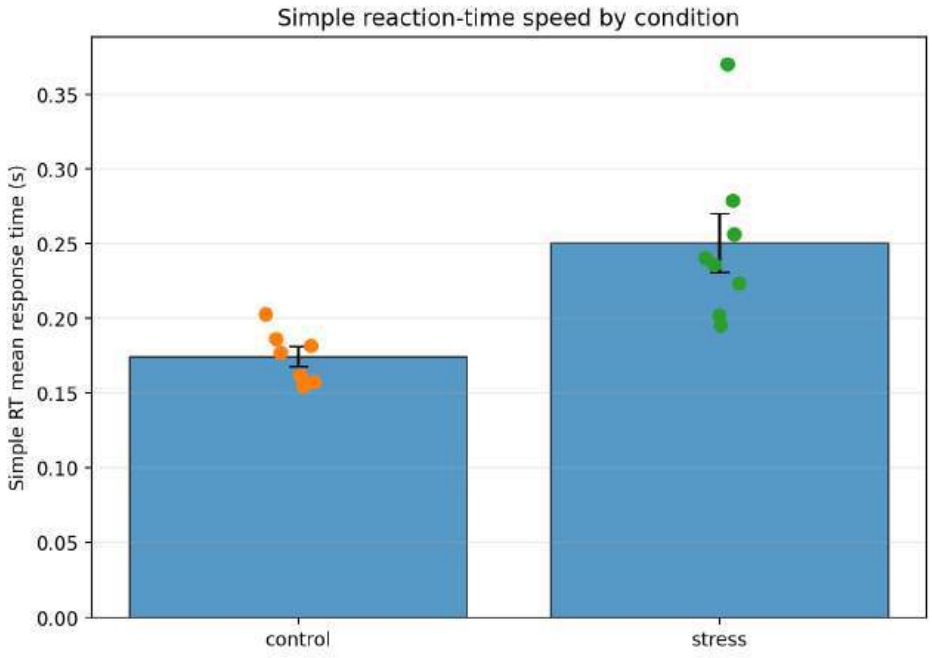


Figure 4, mean reaction time.
Participants exposed to stress had longer simple reaction times than the control group (0.250 s vs. 0.174 s), representing a 30.32% increase. This difference was statistically significant (p = 0.006).

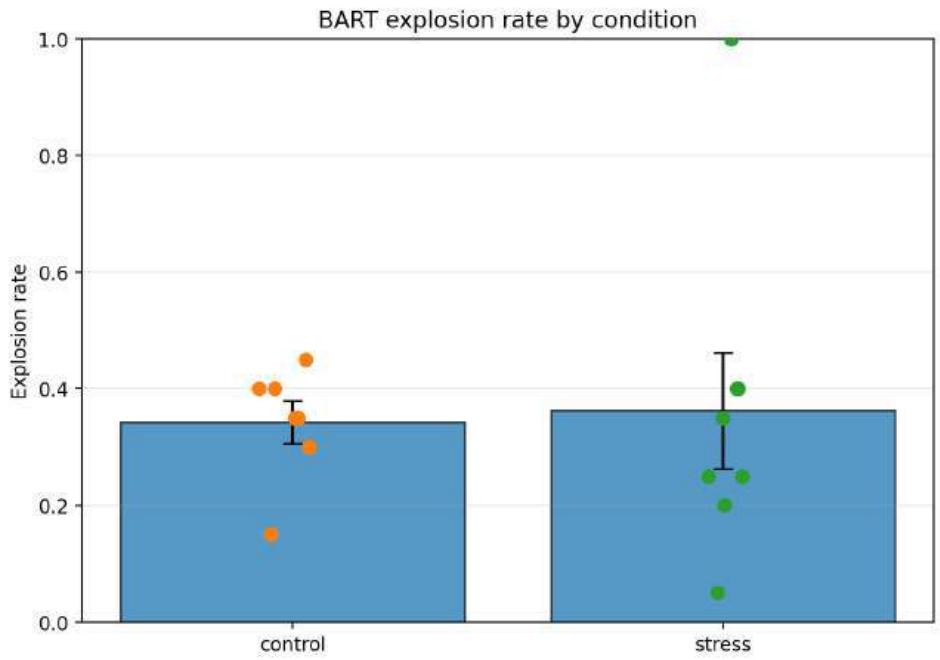


Figure 5, explosion rate. (BART).
The explosion rate did not differ significantly between the stress and control groups (p = 0.858).

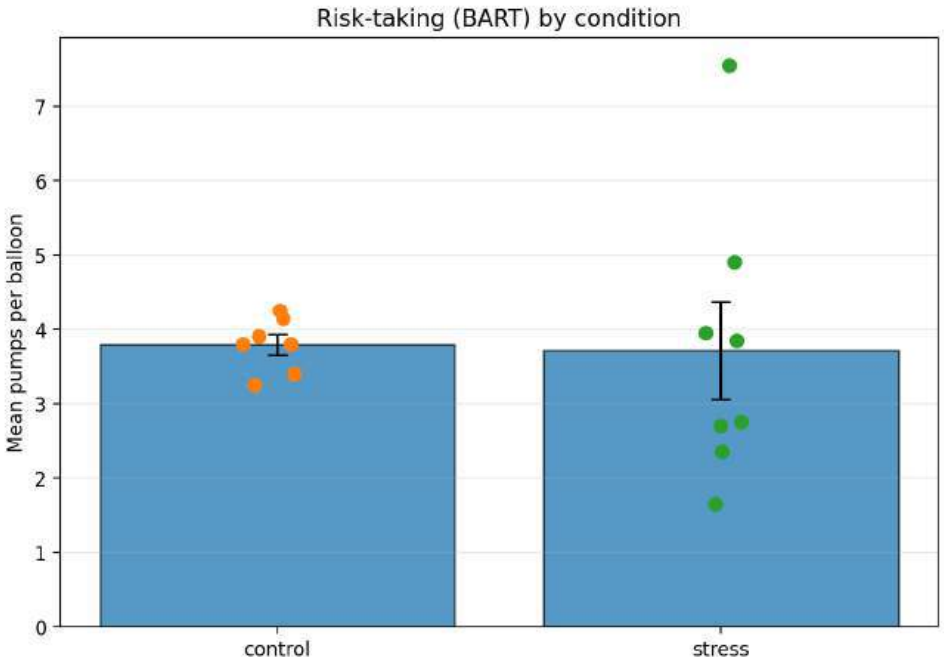


Figure 6, risk-taking by condition.
The mean number of pumps per balloon was almost identical between groups (3.793 vs. 3.713), with only a 2.12% difference. (p = 0.908).

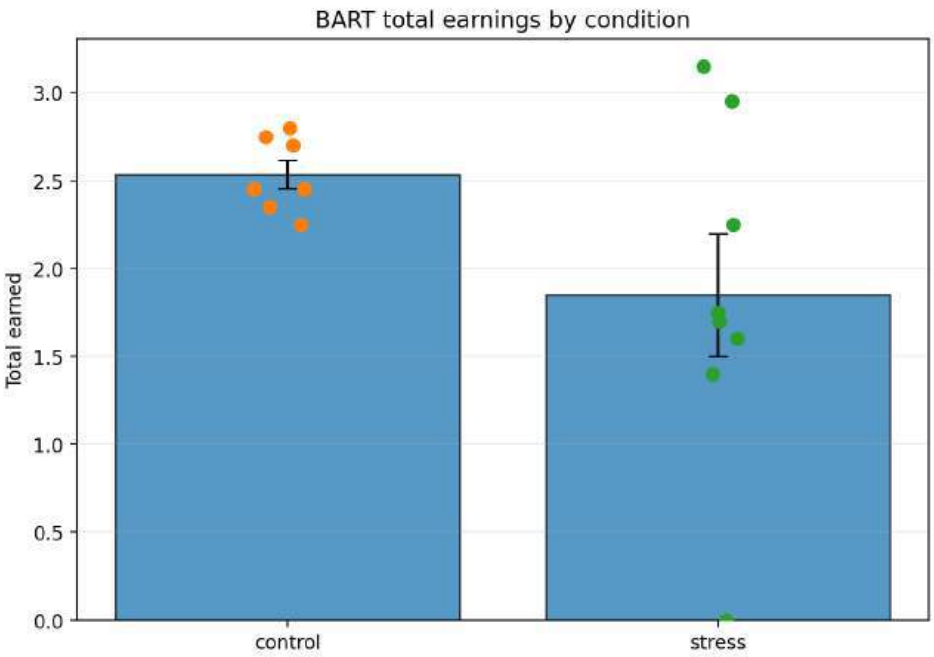


Figure 7, total earnings by conditions.

Participants in the stress group earned less overall than those in the control group (1.850 vs. 2.536), corresponding to a 27.04% reduction in earnings. Although this represents a relatively large numerical difference, it did not reach statistical significance ($p = 0.092$).

Across all measured cognitive outcomes, the average percentage difference between the stress and control groups was approximately 17.15 ± 12.66%.

Discussion

People who underwent the stress induction procedure performed less successfully on working memory tasks than those in the control group. Their accuracy decreased, and overall cognitive performance was lower. These findings support the hypothesis that acute stress temporarily impairs executive functions. One possible explanation is that acute stress increases cortisol levels, which affects the functioning of the prefrontal cortex. Since this brain region plays a central role in working memory, cognitive control, and decision-making, elevated cortisol levels may reduce the ability to maintain and manipulate information effectively.

Reaction time was also negatively affected by stress. Participants exposed to stress demonstrated slower responses during the simple reaction time task, possibly because stress activates the sympathetic nervous system and increases cognitive load. Interestingly, although participants responded faster during some N-back trials, they also made more errors, suggesting a speed–accuracy trade-off. In other words, faster responses may have occurred at the expense of response accuracy.

The effects of acute stress on decision-making were less pronounced. Participants in the stress group earned lower total rewards in the Balloon Analogue Risk Task (BART), while differences in risk-taking behaviour were relatively small and did not reach statistical

significance. These findings suggest that acute stress may influence decision-making; however, its effects appear to be weaker than those observed for working memory and reaction time. Individual differences in stress responses may also have contributed to the variability in performance.

Overall, the present findings are consistent with previous research demonstrating that acute stress negatively affects executive functions and cognitive control (Arnsten, 2009; Shields et al., 2016). Although not all differences reached statistical significance, particularly in the decision-making task, this may be explained by the relatively small sample size, which limited the statistical power of the study.

These findings have important practical implications. Individuals such as students, healthcare professionals, and emergency responders are often required to perform cognitively demanding tasks under stressful conditions. A better understanding of how acute stress influences cognitive performance may contribute to the development of more effective stress-management strategies and cognitive training programs aimed at maintaining performance in high-pressure environments.

Conclusion

I examined how sudden stress affects executive functions in young adults living in Kazakhstan. The results show that stress significantly disrupts working memory and slows reaction time. However, decision-making does not seem to be as strongly affected. These findings are consistent with previous research indicating that stress impairs the brain's ability to manage complex tasks. Although our sample size was small, the results are robust enough to inspire further research, particularly studies with larger samples and additional physiological data.

Work Cited

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. Nature Reviews. Neuroscience, 10(6), 410–422. https://doi.org/10.1038/nrn2648

Baddeley, A. (2003). Working memory: looking back and looking forward. Nature Reviews. Neuroscience, 4(10), 829–839. https://doi.org/10.1038/nrn1201

Dickerson, S. S., & Kemeny, M. E. (2004). Acute Stressors and cortisol Responses: A theoretical integration and synthesis of laboratory research. Psychological Bulletin, 130(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

In brief. (2006). Nature Reviews. Neuroscience, 7(10), 757. https://doi.org/10.1038/nrn2020

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nature Reviews. Neuroscience, 10(6), 434–445. https://doi.org/10.1038/nrn2639

McEwen, B. S. (2007). Physiology and Neurobiology of Stress and Adaptation: Central Role of the Brain. Physiological Reviews, 87(3), 873–904. https://doi.org/10.1152/physrev.00041.2006

Mizuhara, H., Wang, L., Kobayashi, K., & Yamaguchi, Y. (2005). Long-range EEG phase synchronization during an arithmetic task indexes a coherent cortical network simultaneously measured by fMRI. NeuroImage, 27(3), 553–563. https://doi.org/10.1016/j.neuroimage.2005.04.030

Prevot, V., Dehouck, B., Poulain, P., Beauvillain, J., Buée-Scherrer, V., & Bouret, S. (2007). Neuronal–glial–endothelial interactions and cell plasticity in the postnatal hypothalamus: Implications for the neuroendocrine control of reproduction. Psychoneuroendocrinology, 32, S46–S51. https://doi.org/10.1016/j.psyneuen.2007.03.018

Roos, L. E., Knight, E. L., Beauchamp, K. G., Giuliano, R. J., Fisher, P. A., & Berkman, E. T. (2017). Conceptual precision is key in acute stress research: A commentary on Shields, Sazma, & Yonelinas, 2016. Neuroscience & Biobehavioral Reviews, 83, 140–144. https://doi.org/10.1016/j.neubiorev.2017.10.005

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. Neuroscience & Biobehavioral Reviews, 36(4), 1228–1248. https://doi.org/10.1016/j.neubiorev.2012.02.003

Evaluating the Optimal Characteristics for the Addition of Basalt Fibers in Metakaolin-based Geopolymer Materials By Ethan Yao

Abstract

Geopolymers have emerged as a possible substitute for concrete and cementitious materials, particularly due to enhanced strength, energy efficiency, and sustainability. However, geopolymers are deficient in flexural strength, necessitating fiber inclusion to address these shortcomings. Basalt fibers have been conceived as suitable candidates for geopolymer infusion due to their thermal resistance and tensile strength, which could lead to wider applications of these geopolymer materials. This paper compares the results yielded by different mix constituents and proportions to assess their effects on the performance metrics of geopolymer materials. Through analysis of previous research, this study helps determine the optimal fiber and mix characteristics for metakaolin-based geopolymer concrete. This study was conducted as a systematic literature review of papers focusing on basalt fibers and metakaolin-based geopolymer concrete, collecting and comparing data from published experimental papers. Performance metrics of different mixes were analysed to identify how mix constituents and basalt fiber properties affect specimen strength. Fibers with diameters of 20 um and lengths of 12mm provided comparatively higher performance metrics, while the most optimal dosage for basalt fiber was 1.2% by volume for geopolymer mortars and around 0.4% for concretes. This paper aims to outline priorities of future research efforts, ultimately facilitating the commercial and public adoption of geopolymer concrete to promote a more sustainable environment.

Keywords geopolymers, metakaolin, basalt fibers, sodium hydroxide, fiber diameter

Introduction

Concrete and ordinary Portland cement (OPC) production is responsible for a significant proportion of global emissions, with these pollutants accounting for 5-6% of total anthropogenic CO_2 output (Vishojit Bahadur Thapa and Danièle Waldmann, 2026). Geopolymers, a novel group of inorganic materials considered promising substitutes for cementitious building components, have garnered considerable attention due to their enhanced strength and sustainability compared to traditional OPC materials. Geopolymer materials such as concretes and mortars can help reduce CO_2 emissions while also promoting waste recycling and long-term sustainability (Abhishek Sharma et al., 2026). These novel materials are formed from geopolymerization reactions, which solidify inorganic aluminosilicate precursors such as metakaolin and fly ash. Additionally, this process can be accelerated with calcium and heat. The process of geopolymerization can be divided into stages, including destruction (via alkaline solution), polymerization, and stabilization, illustrated by Fig 1 (Xiao Yao et al., 2009). Curation can also be used to harden and strengthen geopolymer materialsHigher temperatures result in increased strength, although excessive heat can cause deformations (Muhd Nurruddin et al., 2018).

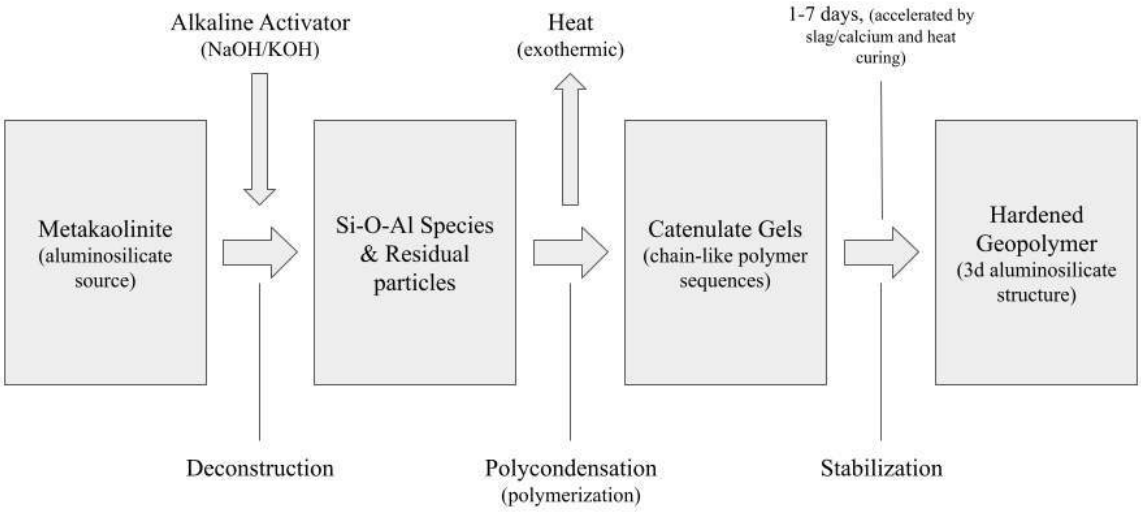


Fig 1: Visualization of the geopolymerization process, including deconstruction, polymerization, stabilization, and the final product of rigid 3d aluminosilicate networks.

Geopolymers can be transformed into cementitious materials by adding aggregates and other compounds. As binders, they can be used as suitable building materials or a pozzolanic paste. Geopolymer mortars can be created through the addition of fine aggregates, binding larger units of building materials (Andrea Kirschner and Harald Harmuth, 2004; Jadambaa Temuujin et al., 2010). Additionally, the inclusion of coarse or structural aggregates can create geopolymer concrete, positioning itself as a possible substitute for OPC (Rafik Abbas et al., 2020). However, geopolymer materials often exhibit low flexural resistance and high brittleness, similar to OPC materials, which restricts their widespread structural applications (Xiao Yao et al., 2009). The addition of fibers to geopolymer materials helps alleviate these deficits by increasing flexural strength and bridging cracks and fractures. These changes could allow geopolymer concrete to become more widely used in infrastructure.

To best expand the uses of geopolymer concrete, it is necessary to assess the characteristics and qualities of different fibers to identify the most effective reinforcement options. Basalt fibers appear particularly promising, largely due to their immense tensile strength and thermal resistance, which could potentially be applied to both underwater and heat-demanding applications (Abhishek Sharma et al., 2026; Mir Ali and M M Vijayalakshmi Natrajan, 2021). These fibers originate from natural igneous basalt rock, which is formed from cooled lava. Some basalt fibers of certain compositions can withstand temperatures of around 1000°C, demonstrating immense thermal stability compared to other fiber types such as bamboo or polypropylene. To curate fibers from rock, basalt is first crushed and heated to around 1500°C in a kiln or furnace without additional chemicals. After heating, the melted basalt is extruded with spinneret technology, producing consistent strands of adjustable diameter with customizable length. Due to the abundance of basalt rock, these fibers prove to be a natural and sustainable form of reinforcement for geopolymer materials (Kunal Singha, 2012). Although the heat

resistance of geopolymer materials is already above average, the addition of basalt to this composite could further enhance its performance under extreme conditions, thereby improving structural integrity and fracture resistance in high-temperature environments.

However, many adjustable characteristics – such as length, dosage, and diameter – also play a crucial role in the strength of basalt fiber and thus the overall structural integrity of the geopolymer binder. For example, longer fibers provide more ductility at extensive cracking scenarios for cementitious materials, while shorter fibers prevent micro-cracks and fractures at a smaller scale (Mingli Cao et al., 2019). Thus, it is necessary to evaluate the optimal characteristics of basalt fiber to ensure maximum physical properties of the geopolymer material, with the ultimate goal of expanding geopolymer applications.

Methodology

This study was conducted as a literature review examining the effects of basalt fibers and different mix proportions on geopolymers. The primary objective of this paper is to determine the optimal mix constituents and basalt fiber characteristics for geopolymer materials. Journal papers published between 2018 and 2026 from academic databases were evaluated to ensure that the collected data were relevant to current research and technology. Studies were included if they used basalt fiber and reported physical properties – such as length and tensile strength – to test flexural and compressive strength metrics. Basalt fiber characteristics, including length, diameter, and dosage, were examined to determine optimal properties for geopolymer reinforcement. Additionally, curation periods of the geopolymer were also assessed.

Data concerning mix constituents and listed mixing procedures were collected to identify trends and relations. Flexural and compressive strength increases were compared to a listed control sample to highlight improvement. These variables were recorded in this comparative study because of their consistent inclusion across papers. Additionally, flexural and compressive strength are primary characteristics evaluated when considering the application of these materials. These performance metrics were compared to the control sample to determine the extent to which the inclusion of a particular basalt fiber sample improved or degraded the strength of the geopolymer. Compressive and flexural values were also compared to the qualities of the basalt fiber to determine a correlation between fiber characteristics and strength.

Results

Table 1 presents the fiber characteristics of studies assessed in this paper. Fiber lengths of 12 mm are the most abundant, while most studies used basalt fibers of 20 μm in diameter. The elastic modulus of the utilized fibers ranges from 88 to 100 GPa

Table 1: Basalt fiber characteristics of studies assessed.

Author, Year	Length (mm)	Diameter (μm)	Nominal tensile strength (MPa)	Elastic Modulus (GPa)	Fiber Density (g/cm^3)
F Şahin et al., 2021	12	20	4100	88	2.73

N Ali et al., 2020	12, 24	13, 20	4100, 4800	89	2.8
N Ali et al., 2018	12	13, 20	4100, 4800	89	2.8
M Uysal et al., 2023	12		4100		2.73
S. Karahüseyin and Ş. Erdoğdu, 2026	12	9-23	4840		2.7
X Zhang et al., 2023	18	15	4300	95	2.67
B. A. Tayeh et al., 2024	10	10	3000	100	2.7
A. A. Arslan et al., 2019	12	20	4100		2.73
A. Celik et al., 2018	12	20	4100		2.73
M. Ziada et al., 2013	12			88	2.73

Table 2 presents the optimal curation time and material type for a wide range of studies with geopolymers. Studies used different testing ages to determine the early- and long-term behavior of geopolymer samples and how performance changes over time. Curation was used to improve the properties of the specimen samples, which were then stored at ambient temperature until the testing day. Most studies investigated mortars.

Table 2: Material type, testing age, and curation characteristics of geopolymer materials with basalt fibers.

Author, Year	Material type	Tested age (days)	Curation time (days)	Curation temperature (°C)
F Şahin et al., 2021	Mortar	7, 28, 56	3	60
N Ali et al., 2020	Mortar	7, 28	3	60
N Ali et al., 2018	Mortar	7, 28	3	60
M Uysal et al., 2023	Mortar	7, 28	1	60
S. Karahüseyin and Ş. Erdoğdu, 2026	Concrete	28	1	60
X Zhang et al., 2023	Binder	7	6	60
B. A. Tayeh et al., 2024	Concrete		1/3	75
A. A. Arslan et al., 2019	Mortar	7, 28	1	105
A. Celik et al., 2018	Mortar	3, 7, 28	3	60
M. Ziada et al., 2013	Mortar	28		

Table 3 presents a variety of stirring techniques utilized for mixing geopolymers and basalt fibers. Many studies incorporated similar methods to mixing procedures for cementitious materials, and some used a vibration table after raw mechanical stirring.

Table 3: Mixing techniques for geopolymer materials with basalt fibers

Author, Year	Mixing Technique
F Şahin et al., 2021	Mechanical stirrer + vibration
N Ali et al., 2020	Stirrer Drill + vibration
N Ali et al., 2018	Mechanical stirrer + vibration
M Uysal et al., 2023	Not explicitly stated
S. Karahüseyin and Ş. Erdoğdu, 2026	Mechanical Mixer
X Zhang et al., 2023	Full-automatic planetary cement mixer + vibration table
B. A. Tayeh et al., 2024	Mechanically Adjustable mixer
A. A. Arslan et al., 2019	Stirrer Drill + vibration
A. Celik et al., 2018	Not explicitly stated
M. Ziada et al., 2013	Blender + vibration table

Table 4 presents mix constituents and their mass proportional to the total. Metakaolin, red mud, pumice, fine aggregates, waste marble powder, basalt powder, limestone, sodium silicate solution, sodium hydroxide, and sodium silicate were abbreviated to MK, RM, PM, FA, BP, LS, SSS, SH, and SS, respectively. The “Other” portion of the table consisted of a 1:1 mix of water and plasticizer. Most studies employed a sodium hydroxide-to-sodium silicate ratio of 1:2. Additionally, multiple studies used approximately 22% metakaolin and 3% slag for mortars. In contrast, studies that tested concrete used a relatively lower amount of metakaolin. Aggregate percentages for most mortars were around 54%, while concretes used small amounts of aggregates.

Table 4: Mix proportions of geopolymer materials with basalt fibers (% of total mass). Metakaolin, red mud, pumice, fine aggregates, waste marble powder, basalt powder, limestone, sodium silicate solution, sodium hydroxide, and sodium silicate were abbreviated to MK, RM, PM, FA, BP, LS, SSS, SH, and SS, respectively

Author, Year	MK	Slag	Zeolite	RM	PM	FA	WMP	BP	LS	SSS	SH	Other
F Şahin et al., 2021	22.80 %	3.00%				51.30 %				15.20 %	7.60%	
N Ali et al., 2020	21.60 %	2.90%				54.00 %				14.40 %	7.20%	
N Ali et al.,	21.60	2.90%				54.00				14.40	7.20%	

2018	%					%				%		
M Uysal et al., 2023	10.10 %	0.90%		10.10 %		38.00 %		12.70 %		17.60 %	8.80%	1.80%
S. Karahüseyin and Ş. Erdoğdu, 2026	5.60%	16.90 %			17.30 %				42.80 %	11.60 %	5.80%	
X Zhang et al., 2023	46.10 %									48.31 %	5.59%	
B. A. Tayeh et al., 2024	6.50%	30.50 %	6.50%			6.50%	24.00 %			18.50 %	7.40%	
A. A. Arslan et al., 2019	21.60 %	2.90%				54.00 %				14.40 %	7.20%	
A. Celik et al., 2018	21.60 %	2.90%				54.00 %				14.40 %	7.20%	
M. Ziada et al., 2013	14.09 %	12.54 %				51.83 %				14.36 %	7.18%	

Table 5 illustrates the basalt fiber percentage based on the mass of the geopolymer material, though some studies chose to assess the dosage percentage at the binder stage. Additionally, studies with percentage representations not equal to the total mass were converted to ensure consistent comparison. Dosage percentages range from 0.25% to 2.4% of total mass, not exceeding certain values, potentially due to workability issues.

Table 5: Fiber dosage percentage based on mass for geopolymer materials with basalt fiber

Author, Year	Dosage % of representation mass	Dosage Representation	Dosage % of total mass
N Ali et al., 2018	1%	Binder Mass	0.25%
X Zhang et al., 2023	1, 3, 5%	Binder Mass	0.4, 1.4, 2.3%
B. A. Tayeh et al., 2024	1, 2%	Binder Mass	0.4, 0.8%
A. A. Arslan et al., 2019	0.8, 1.6, 2.4%	Total Mass	0.8, 1.6, 2.4%

Table 6 illustrates the volume of basalt fiber compared to the total volume of the geopolymer material. Percentages range from 0.1% to 1.2%, with 1% and 1.2% as the most common maximums.

Table 6: Fiber dosage percentage based on volume for geopolymer materials with basalt fiber

Author, Year	Dosage % of total volume.
F Şahin et al., 2021	0.4%, 0.8%, 1.2%

N Ali et al., 2020	1%
M Uysal et al., 2023	0.4, 0.8, 1.2%
S. Karahüseyin and Ş. Erdoğdu, 2026	0.1, 0.2, 0.3, 0.4, 0.5, 0.6%
A. Celik et al., 2018	0.4, 0.8,1.2%
M. Ziada et al., 2013	0.5, 0.75, 1%

Table 7 compares the compressive strength of geopolymer materials infused with different fiber dosages to the control sample, with masses standardized to total mass to ensure consistency. The value corresponding to the ambient temperature storage duration that yielded the highest compressive strength is used in this table. While the study that tested mortars benefited from larger quantities of basalt fiber, other material types had a higher compressive strength at lower amounts of basalt fiber. For example, non-mortar specimens containing 0-0.4% basalt fiber led to strengths of 60.3-67.1 MPa. The best percentage of basalt fiber infusion for each respective study is marked with an asterisk.

Table 7: The compressive strength (CS) of basalt fiber-infused geopolymer concrete based on total-mass dosage percentages.

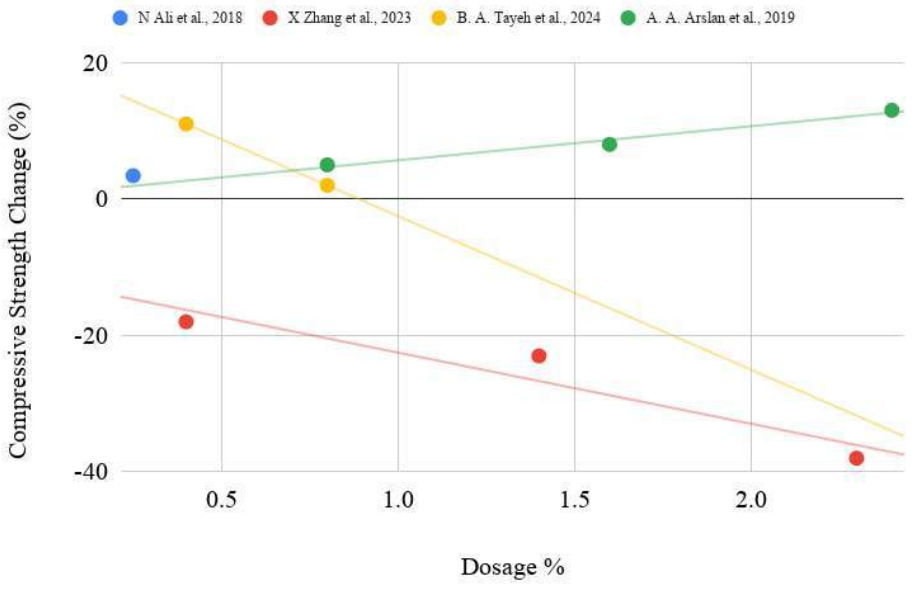
		Dosage Percentage						
Author, Year	Control CS (MPa)	0.25%	0.4%	0.8%	1.4%	1.6%	2.3%	2.4%
N Ali et al., 2018	48.37	50*						
X Zhang et al., 2023	60.3*		41.2		38.4		30.9	
B. A. Tayeh et al., 2024	55		67.1*	61.2				
A. A. Arslan et al., 2019	63			63		65		68*

Table 8 compares the compressive strength of geopolymer materials infused with varying fiber dosages to the control sample. The value corresponding to the ambient temperature storage duration that yielded the highest compressive strength is used in this table. The dosage percentages of 1% and 1.2% provided the highest compressive strength values, creating specimens whose strength ranges from 34.3 to 76 MPa. Most mortars achieved their greatest flexural strength at higher dosage percentages, while concrete samples attained the most meaningful improvement from lower dosage percentages. The best compressive value of basalt fiber infusion for each respective study is marked with an asterisk.

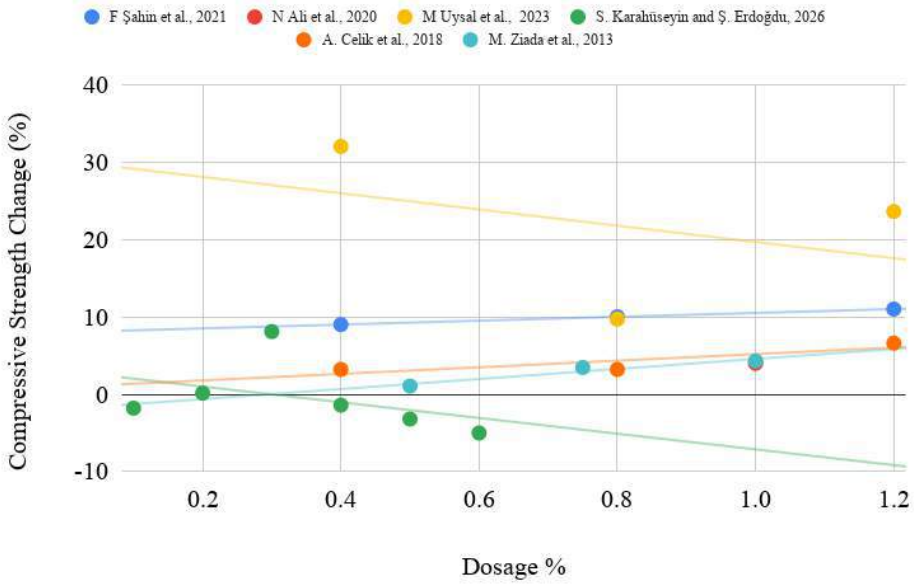
Table 8: The compressive strength of basalt fiber-infused geopolymer concrete based on total-volume dosage percentages.

		Dosage Percentage									
Author, Year	Control CS (MPa)	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.75%	0.8%	1%	1.2%
F Şahin et al., 2021	68				74				74		76*
N Ali et al., 2020	51									53*	
M Uysal et al., 2023	28.41				38.4*				31.8		34.3
S. Karahüseyin and Ş. Erdoğdu, 2026	22	21.6	25*	23.8	23.5	22.7	20.9				
A. Celik et al., 2018	63				63				63		65*
M. Ziada et al., 2013	58					58.6		60		60.5*	

Fig 2a compares the compressive strength of mass-based basalt fiber geopolymer material samples with dosage percentage. Most studies that did not test mortars saw decreased compressive strength with further addition of basalt fiber. Fig 2b compares the compressive strength of volume-based basalt fiber geopolymer material samples with dosage percentage. Most studies that tested mortars benefited from increasing basalt fiber, while other materials, such as concretes, had an opposite effect. A scatter plot was used to visualize trends in the data.



(a)



(b)

Fig 2: (a) Total-Mass-based and (b) volume-based basalt fiber dosage percentage and compressive strength change compared to the control sample.

Table 9 displays the flexural strengths of various basalt fiber-infused geopolymer specimens compared to the control sample. The value corresponding to the ambient temperature storage duration that yielded the highest flexural strength is used in this table. Similar to Table 7, the study that tested mortars achieved its highest flexural strength at the highest tested dosage percentage, while other studies saw higher performance metrics at lower dosages. The highest flexural strength for a mortar was provided by 2.4%, while lower values around 0.4% performed well for the binder and concrete, ranging from 7.52-10.93 MPa of flexural resistance. The best samples for each respective study are marked with an asterisk.

Table 9: The flexural strength (FS) of basalt fiber-infused geopolymer concrete based on total-mass dosage percentages.

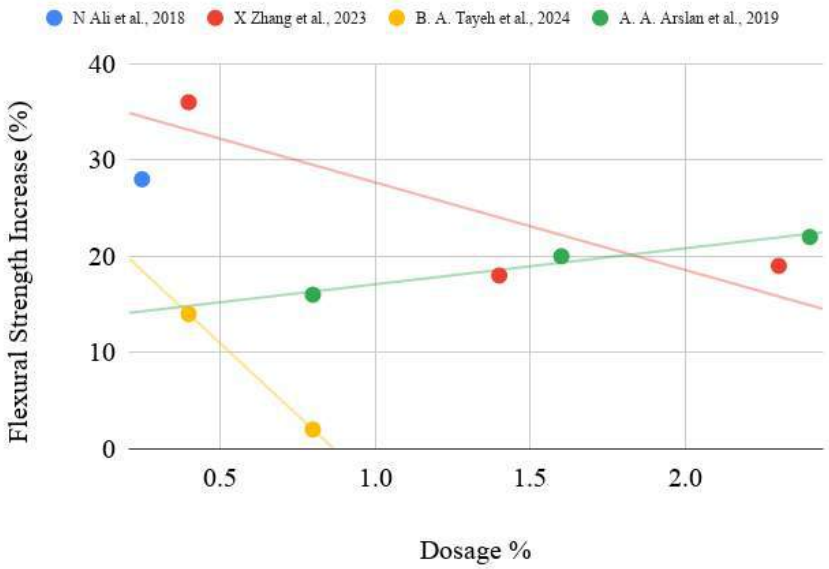
		Dosage Percentage						
Author, Year	Control FS (MPa)	0.25%	0.4%	0.8%	1.4%	1.6%	2.3%	2.4%
N Ali et al., 2018	8.9	11.4*						
X Zhang et al., 2023	8.06		10.93*		9.54		9.63	
B. A. Tayeh et al., 2024	6.61		7.52*	6.72				
A. A. Arslan et al., 2019	10.3			11.9		12.4		12.6*

Table 10 displays the compressive strengths of various basalt fiber-infused geopolymer specimens compared to the control sample. The value corresponding to the ambient temperature storage duration that yielded the highest flexural strength is used in this table. Most studies achieved the best flexural strength at the highest dosage percentage, while the study that tested geopolymer concrete saw its highest flexural strength at a middle dosage percentage. Mortar specimens that included around 1-1.2% basalt fiber saw flexural strengths ranging from 10.6-12.6 MPa. The best samples for each respective study are marked with an asterisk.

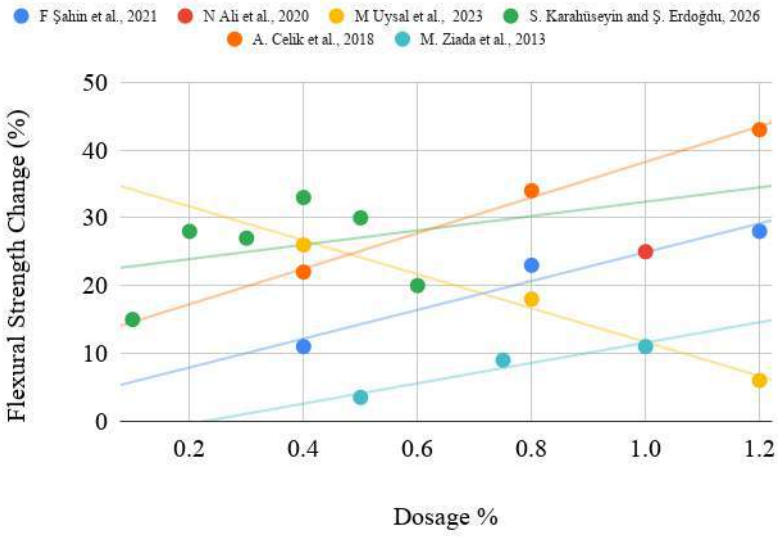
Table 10: The flexural strength of basalt fiber-infused geopolymer concrete based on total-volume dosage percentages.

		Dosage Percentage									
Author, Year	Control CS (MPa)	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.75%	0.8%	1%	1.2%
F Şahin et al., 2021	8.8				9.8				10.8		11.3*
N Ali et al., 2020	9.4									11.8*	
M Uysal et al., 2023	10				12.6*				11.8		10.6
S. Karahüseyin and Ş. Erdoğdu, 2026	1.63	1.9	2.4	2.1	2.8*	3.6	2.0				
A. Celik et al., 2018	8.8				10.7				11.8		12.6*
M. Ziada et al., 2013	10.5					10.9		11.4		11.7*	

Fig 3a compares the flexural strength of mass-based basalt fiber geopolymer material samples with dosage percentage, converted to total mass if expressed in binder mass. Arslan et al.'s study (2019) saw an increased relationship between flexural strength and basalt fiber dosage, while others displayed a decreasing trend or had limited data. Fig 3b compares the flexural strength of volume-based basalt fiber geopolymer material samples with dosage percentage. A scatter plot was used to determine trends in the data.



(a)



(b)

Fig 3: (a) Total-Mass-based and (b) volume-based basalt fiber dosage percentage and flexural strength change compared to the control sample.

Discussion

These studies demonstrate that basalt fiber generally has a positive effect on the compressive strength of geopolymer materials, although the magnitude of improvement at

different fiber percentages is heavily dictated by material type. This improvement is likely due to the fibers' role in distributing stress and delaying crack initiation, which increases both compressive and flexural strength. The highest compressive strength for a geopolymer sample was reported by F. Sahin et al. (2021), producing a mortar with a strength of 76 MPa by utilizing a dosage percentage of 1.2% by volume. For mass-based dosage percentages, the geopolymer mortar tested by A. A. Arslan et al. (2019) yielded a compressive strength of 68 MPa at 2.4% basalt fiber. Additionally, in studies from F Şahin et al. (2021), A. A. Arslan et al. (2019), and A. Celik et al. (2018) – whose mix proportions are similar– the highest compressive strength results came from samples with the highest basalt fiber dosage percentage.

However, for geopolymer binders and concrete, excessive fiber volumes degrade performance at lower dosage percentages than mortars. For example, X Zhang et al. (2023) saw a sharp decrease from 60.3 MPa control to 30.9 MPa when fiber content was increased to 2.3% due to increased pore volume and decreased density from lower workability. In these types of geopolymer materials, the presence of coarse aggregates could create challenges of fiber dispersion and workability, reducing basalt fiber's effectiveness. Overall, optimal performance metrics are achieved at basalt fiber dosages of 1-1.2% and 0.2-0.4% for mortars and concretes/binders, respectively. Below these limits, fibers successfully densify the geopolymer's internal structure by filling pores; however, excessive fiber amounts lead to workability losses and air entrapment, which compromise matrix compaction and thus structural integrity.

These studies also demonstrate that basalt fiber significantly enhances the flexural strength of geopolymer materials, particularly in mortars. High flexural strengths included A. Arslan et. al. (2019) at 2.6% basalt fiber by total mass, M. Uysal et. al.(2021) with 0.4% basalt fiber by volume, and A. Celik et al. (2018) at 1.2% basalt fiber by volume, all yielding flexural strengths of 12.6 MPa. Specimens are strongest at the highest dosage percentage (1.2% by volume and 2.4% by mass) in studies by F Şahin et al. (2021), A. A. Arslan et al. (2019), and A. Celik et al. (2018), suggesting a positive relationship between basalt fiber content and geopolymer mortar strength within the investigated ranges. However, the effect of dosage percentage is highly dependent on material type. For mortars, multiple studies saw a consistently increasing trend proportional to the amount of basalt fiber added. In contrast, studies by X. Zhang et al. (2023) and B. A. Tayeh et al. (2024), which tested geopolymer binders and concretes, respectively, both reported a decrease in flexural strength improvement when excessive basalt fiber was added. The decreases could be a result of internal clumping and air entrainment due to low workability from excessive basalt fiber. Assessing sample mixability when infusing basalt fibers is especially important for geopolymer concretes, as coarse aggregates, such as limestone, can absorb water and create internal friction that decreases workability (Mingli Cao et al., 2019). Although many studies display an increasing trend in flexural strength with the addition of basalt fibers, Fig 3a and 3b illustrate that the increase in strength is proportional to the basalt fiber dosage plateaus, but more likely drops off at higher dosage percentages after optimal dosage. A demonstration of this decline in strength is most profoundly shown by S. Karahüseyin and Ş. Erdoğdu's (2026) concrete sample, which contains a

positive trendline but a decrease in strength after 0.4% basalt fiber dosage. Thus, there exists an optimal dosage percentage of basalt fiber inclusion where workability issues do not overwhelm flexural strength improvements from basalt fiber inclusion

The geopolymer's 3d aluminosilicate structure is also shown to evolve as ongoing geopolymerization and densification continue; for most studies, 28 days of storage at an ambient temperature lead to higher performance metrics than 7 days. For curation, most studies chose to cure the specimens at a temperature of 60° Celsius for 1-3 days. P. Rovnaník's (2010) study on the strengths of metakaolin-based geopolymer mortars corresponds with the commonality of this temperature and curation period, with his paper noting that the highest physical properties came from samples cured at 60° Celsius for 1 day, which demonstrates the effectiveness of curation at similar temperatures.

Mixing proportions of different substances and the characteristics of basalt fibers directly impact the performance metrics of geopolymer materials. For geopolymer mortars, 20% metakaolin and 3% slag were commonly used among studies; mortars F Şahin et al. (2021), N Ali et al. (2020), N Ali et al. (2018), A. A. Arslan et al. (2019), and A. Celik et al. (2018) all contain similar metakaolin and slag amounts while also providing relatively high compressive and flexural strength. However, geopolymer concrete samples use significantly more slag than metakaolin; the addition of calcium can contribute to early strength development and solidification, which helps account for the coarser aggregates and provides early strength by reducing setting time (Yiren Wang et al., 2024). The majority of studies used a sodium hydroxide to sodium silicate ratio of 1:2, with around 14-16% sodium silicate and 7-8% sodium hydroxide. A. Albidah et. al. note that geopolymer concretes with sodium hydroxide: sodium silicate solutions of around 1:2 provide not only increased workability but also high compressive strength (Abdulrahman Albidah et al., 2021). Mixability is especially important in this case, as the addition of basalt fibers already creates workability issues and can lead to internal defects and air entrainment.

Furthermore, the physical properties of basalt fibers directly affect the performance metrics of the geopolymer material. Studies by F Şahin et al. (2021), N Ali et al. (2020), N Ali et al. (2018), A. A. Arslan et al. (2019), and A. Celik et al. (2018) all included basalt fibers with a diameter of 20μm and produced mortars with greater flexural and compressive strength compared to other studies. This could be a result of higher load-bearing properties from higher rigidity from thickness, although fibers with excessive thickness can contain more internal flaws and imperfections (Yixuan Ma et al., 2025). Additionally, 12mm and 20 μm are the most commonly utilized lengths and diameters for basalt fibers, although this may not fully represent the most optimal value. At longer lengths, the mixability of the binder decreases, causing additional air entrainment and unnecessary pore formation that ultimately reduces the overall strength of the sample. Excessively long fibers can also clump and entangle themselves, reducing their potential to increase the strength of the geopolymer specimen. However, the infusion of longer fibers in cementitious materials can also help bridge unavoidable pores and prevent fracturing, which can result in higher flexural and compressive resistance (Abdulrahman Albidah

et al., 2021). To test different types of basalt fiber, N Ali et. al. (2020) used basalt fibers of 12mm length and 13μm diameter (SBF) in addition to 24mm length and 20μm diameter (TBF). Although the compressive strengths of geopolymer mortars infused with SBF and TBF were largely similar, the flexural strength of mortars containing TBF was significantly greater than that of SBF. This suggests that longer basalt fibers with higher diameters improve flexural strength at a larger magnitude, although additional mixing procedures, such as vibration, are recommended to release trapped air caused by lower workability. Additionally, M Heweidak et al. (2022) tested multiple geopolymer samples containing single-length and hybrid-length basalt fibers, concluding that specimens containing a combination of 30mm and 12mm basalt outperformed samples with a single fiber length. Thus, the incorporation of a hybrid fiber composition with both tall and short basalt fibers can result in advantageous effects, as a combination of lengths could help reduce fractures at a wide range of magnitudes while also ensuring manageable workability.

Conclusion

This study analysed previous research to determine the optimal mix constituents and basalt fiber characteristics for geopolymer materials to promote their usage and thus a more sustainable environment. The inclusion of basalt fiber was ultimately shown to significantly improve the mechanical properties of geopolymer materials if included in optimal quantities. If basalt fiber inclusion exceeds a certain percentage of around 0.4% for concrete/binders and 1-1.2% for mortars, compressive/tensile performance begins to decrease, likely due to low workability and increased air entrainment, leading to higher porosity and thus brittleness. Additionally, mortars saw consistent mix constituent proportions of around 1.2% basalt fiber by mass with 22% metakaolin, 3% slag, 54% fine aggregates, 14% sodium silicate solution, and 7% sodium hydroxide. There exists a peak in strength increase where the basalt fibers can bridge cracks while also ensuring manageable workability, which varies greatly with mix constituents and proportions and thus requires minor testing. Fiber characteristics can influence its mechanical properties and further increase its effectiveness in applications. A length of 12mm was most common among studies, although longer fibers with higher diameter or a hybrid fiber composition of both tall and short basalt fibers could result in advantageous effects

Acknowledgements

I would like to thank my mentor, Nate Buettner, for guiding me through research and writing.

Works Cited

Abbas, Rafik, et al. "Preparation of Geopolymer Concrete Using Egyptian Kaolin Clay and the Study of Its Environmental Effects and Economic Cost." Clean Technologies and Environmental Policy, vol. 22, no. 3, 2020, pp. 669–87. Accessed 12 July 2026. doi.org/10.1007/s10098-020-01811-4.

Albidah, Abdulrahman, et al. "Characteristics of Metakaolin-Based Geopolymer Concrete for Different Mix Design Parameters." Journal of Materials Research and Technology, vol. 10, 2021, pp. 84–98. Accessed 12 July 2026. doi.org/10.1016/j.jmrt.2020.11.104.

Ali, Mir Firasath, and M. M. Vijayalakshmi Natrajan. "A Review of Geopolymer Composite Thermal Properties." IOP Conference Series: Earth and Environmental Science, vol. 822, no. 1, 2021, article 012051. Accessed 12 July 2026. doi.org/10.1088/1755-1315/822/1/012051.

Ali, Nawar, et al. "Effect of Using Colemanite and Basalt Fiber on the Mechanical Properties of Metakaolin-Based Geopolymer Mortars." Journal of Sustainable Construction Materials and Technologies, vol. 3, no. 2, 2018, pp. 235–41. Accessed 12 July 2026. doi.org/10.29187/jscmt.2018.25.

Ali, Nawar, et al. "Evaluation of the 12–24 mm Basalt Fibers and Boron Waste on Reinforced Metakaolin-Based Geopolymer." Construction and Building Materials, vol. 251, 2020, article 118976. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2020.118976.

Arslan, Ahmet Ali, et al. "Influence of Wetting-Drying Curing System on the Performance of Fiber Reinforced Metakaolin-Based Geopolymer Composites." Construction and Building Materials, vol. 225, 2019, pp. 909–26. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2019.07.235.

Blanco, Ignazio, and Michelina Catauro. "Geopolymers—Design, Preparation, and Applications." Polymers, vol. 14, no. 5, 2022, article 853. Accessed 12 July 2026. doi.org/10.3390/polym14050853.

Canpolat, Orhan, et al. "Effect of Using Colemanite and Basalt Fiber on the Mechanical Properties of Metakaolin-Based Geopolymer Mortars." Journal of Sustainable Construction Materials and Technologies, vol. 3, no. 2, 2018, article 25. Accessed 12 July 2026. doi.org/10.29187/jscmt.2018.25.

Cao, Mingli, et al. "Effect of Macro-, Micro-, and Nano-Calcium Carbonate on Properties of Cementitious Composites—A Review." Materials, vol. 12, no. 5, 2019, article 781. Accessed 12 July 2026. doi.org/10.3390/ma12050781.

Celik, Abdullah, et al. "High-Temperature Behavior and Mechanical Characteristics of Boron Waste Additive Metakaolin-Based Geopolymer Composites Reinforced with Synthetic Fibers." Construction and Building Materials, vol. 187, 2018, pp. 1190–203. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2018.08.062.

Heweidak, Mohamed, Bidur Kafle, and Riyadh Al-Ameri. "Influence of Hybrid Basalt Fibres' Length on Fresh and Mechanical Properties of Self-Compacted Ambient-Cured

Geopolymer Concrete." Journal of Composites Science, vol. 6, no. 10, 2022, article 292. Accessed 18 July 2026. https://doi.org/10.3390/jcs6100292.

Karahüseyin, Sümeyye, and Şakir Erdoğdu. "Assessment of the Effect of Basalt Fiber on the Mechanical Behavior and Thermal Conductivity Properties of Structural Lightweight Geopolymer Concrete." Innovative Infrastructure Solutions, vol. 11, no. 2, 2026, article 2471. Accessed 12 July 2026. doi.org/10.1007/s41062-025-02471-3.

Kirschner, Andrea V., and Harald Harmuth. "Investigation of Geopolymer Binders with Respect to Their Application for Building Materials." Ceramics–Silikáty, vol. 48, no. 3, 2004, pp. 117-124. Accessed 12 July 2026. irsm.cas.cz/materialy/cs_content/2004/Kirschner_CS_2004_0000.pdf.

Ma, Yixuan, et al. "From Rock to Fiber: The Mechanical Properties of Continuous Rock Fibers." Natural Resources Research, vol. 34, no. 4, 2025, pp. 1953–85. Accessed 12 July 2026. doi.org/10.1007/s11053-025-10483-0.

Nurruddin, et al. "Methods of Curing Geopolymer Concrete: A Review." International Journal of Advanced and Applied Sciences, vol. 5, no. 1, 2018, pp. 31–36. Accessed 12 July 2026. doi.org/10.21833/ijaas.2018.01.005.

Rovnaník, Pavel. "Effect of Curing Temperature on the Development of Hard Structure of Metakaolin-Based Geopolymer." Construction and Building Materials, vol. 24, no. 7, 2010, pp. 1176–83. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2009.12.023.

Şahin, Furkan, et al. "Effect of Basalt Fiber on Metakaolin-Based Geopolymer Mortars Containing RILEM, Basalt, and Recycled Waste Concrete Aggregates." Construction and Building Materials, vol. 301, 2021, article 124113. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2021.124113.

Sharma, Abhishek, et al. "Potential of Geopolymer Concrete as Substitution for Conventional Concrete: A Review." Materials Today: Proceedings, vol. 57, 2022, pp. 1539–45. Accessed 12 July 2026. doi.org/10.1016/j.matpr.2021.12.159.

Singha, Kunal. "A Short Review on Basalt Fiber." International Journal of Textile Science, vol. 1, no. 4, 2012, pp. 19–28. Allendale Fiber Tech, allendalefibertech.com/wp-content/uploads/2023/02/Basalt-Technical-Fibers.pdf. Accessed 12 July 2026.

Tayeh, Bassam A., et al. "Effect of Hybrid Fibers and High Temperatures on the Properties of Geopolymer Composites Based on Slag, Metakaolin, and Natural Zeolite." Construction and Building Materials, vol. 451, 2024, article 138898. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2024.138898.

Temuujin, Jadambaa, et al. "Preparation and Characterisation of Fly Ash Based Geopolymer Mortars." Construction and Building Materials, vol. 24, no. 10, 2010, pp. 1906–10. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2010.04.012.

Thapa, Vishojit Bahadur, and Danièle Waldmann. "A Short Review on Alkali-Activated Binders and Geopolymer Binders." Vielfalt im Massivbau - Festschrift zum 65. Geburtstag von

Prof. Dr. Ing. Jürgen Schnell, Ernst & Sohn, 2018, pp. 576-591. University of Luxembourg, accessed 12 July 2026. orbilu.uni.lu/handle/10993/35284.

Uysal, Mücteba, et al. "The Effect of Various Fibers on the Red Mud Additive Sustainable Geopolymer Composites." Construction and Building Materials, vol. 363, 2023, article 129864. Accessed 12 July 2026. doi.org/10.1016/j.conbuildmat.2022.129864.

Wang, Yiren, et al. "Effects of Different Calcium Sources on Mechanical Properties of Metakaolin Geopolymers." Materials, vol. 17, no. 9, 2024, article 2087. Accessed 12 July 2026. doi.org/10.3390/ma17092087.

Yao, Xiao, et al. "Geopolymerization Process of Alkali-Metakaolinite Characterized by Isothermal Calorimetry." Thermochimica Acta, vol. 493, no. 1–2, 2009, pp. 49–54. Accessed 12 July 2026. doi.org/10.1016/j.tca.2009.04.002.

Zhang, Xinli, et al. "Bamboo, Basalt, Glass, and Polypropylene Fiber-Reinforced Metakaolin Based Geopolymers: A Comparative Study." European Journal of Wood and Wood Products, vol. 81, no. 6, 2023, pp. 1481–94. Accessed 12 July 2026. doi.org/10.1007/s00107-023-01960-5.

Ziada, Mahmoud, et al. "Influence of Various Fibers on the Physico-Mechanical Properties of a Sustainable Geopolymer Mortar Based on Metakaolin and Slag." Engineering Science and Technology, an International Journal, vol. 46, 2023, article 101501. Accessed 12 July 2026. doi.org/10.1016/j.jestch.2023.101501.

A Theoretical Framework for the Use of Vedic Mathematics in Children with Learning Disabilities and Hearing Loss: Potential Cognitive Mechanisms and Future Research Directions By Aariana Gawande

Abstract

Children with learning disabilities and hearing loss often encounter significant barriers to mathematical achievement, particularly in tasks demanding working memory, processing speed, and executive function. Vedic Mathematics - an ancient Indian system of mental computation based on sixteen sutras - offers a pattern-based, cognitively efficient, and visually oriented approach to arithmetic that may address many of these deficits. This paper develops a theoretical framework examining the potential cognitive mechanisms through which Vedic Mathematics instruction may benefit children with dyscalculia, dyslexia, attention-related disorders, and hearing loss. Drawing on Cognitive Load Theory, the Working Memory Model, Information Processing Theory, Constructivist Learning Theory, and Dual Coding Theory, we propose a conceptual pathway from Vedic Mathematics instruction through reduced cognitive load and improved working memory efficiency to enhanced arithmetic fluency, academic confidence, and reduced math anxiety. I have given particular attention to the unique cognitive profile of children with hearing loss, for whom the visual and non-auditory nature of Vedic techniques may offer a compelling compensatory pathway. I have identified which learning populations are most likely to benefit, highlight current gaps in the empirical literature, and propose specific directions for future experimental research.

Keywords Vedic Mathematics, learning disabilities, dyscalculia, ADHD, hearing loss, cognitive load theory, working memory, mathematical cognition, educational neuroscience, deaf education

Introduction

Mathematical learning disabilities affect approximately 5–8% of school-age children worldwide, presenting persistent challenges in numerical reasoning, arithmetic fluency, and problem-solving (Butterworth et al., 2011). In parallel, children with hearing loss - estimated at 2–3 per 1,000 newborns globally - face compounding barriers to mathematical achievement arising from communication difficulties, increased cognitive load from listening effort, and reduced access to the language-dependent instruction that characterizes most conventional mathematics classrooms (Marschark et al., 2015).

Vedic Mathematics, a system of mental arithmetic and algebraic computation derived from ancient Indian scriptures and systematized by Bharati Krishna Tirthaji in the early twentieth century, has attracted growing interest as an alternative pedagogical framework (Tirthaji, 1965). Based on sixteen sutras, or word formulae, the system emphasizes pattern recognition, mental shortcuts, and stepwise algorithms designed to simplify complex calculations. Its visual, non-sequential, and pattern-based character may align particularly well with the learning profiles of children who struggle in conventionally auditory, language-heavy mathematical instruction.

However, empirical evidence for the efficacy of Vedic Mathematics in populations with learning disabilities or hearing loss remains limited. No randomized controlled trials have been conducted in these populations, and most existing studies focus on typically developing learners. This paper addresses that gap by developing a rigorous theoretical framework, mapping specific population profiles onto Vedic Mathematics components and proposing the cognitive mechanisms through which instruction may produce measurable benefit.

Study Objective

The primary objective of this paper is to develop a theoretical framework explaining how Vedic Mathematics instruction may improve mathematical learning in children with learning disabilities and hearing loss, based on a structured review of current educational, cognitive, and neuroscience literature.

Research Questions

This paper addresses the following research questions:

1. What cognitive processes are most commonly impaired in children with learning disabilities and hearing loss relevant to mathematical performance?
2. Which components of Vedic Mathematics engage these cognitive processes?
3. Is there a theoretical basis for expecting improved mathematical performance through Vedic Mathematics instruction in these populations?
4. Which subgroups are most likely to benefit, and under what instructional conditions?

Literature Review

Cognitive Deficits in Learning Disabilities

Learning disabilities encompass a heterogeneous group of neurodevelopmental conditions that interfere with the acquisition of academic skills. In the domain of mathematics, dyscalculia is the most extensively studied condition, characterized by persistent difficulties in number sense, arithmetic fact retrieval, and procedural computation in the absence of general intellectual impairment (American Psychiatric Association, 2013). Prevalence estimates range from 3% to 7% of the school-age population (Shalev, 2004).

Working memory deficits are among the most consistently documented cognitive features of mathematical learning disabilities. Geary (2004) demonstrated that children with math difficulties show reduced capacity in the phonological loop and central executive components of working memory, impairing their ability to hold interim results during multi-step calculations. Processing speed deficits further compound these difficulties, slowing arithmetic fact retrieval and reducing the automaticity of foundational computations (Bull & Johnston, 1997). Executive function impairments - including deficits in planning, cognitive flexibility, and inhibitory control - are prominent in children with ADHD and dyscalculia, affecting their ability to apply multi-step problem-solving strategies (Toll et al., 2011). Additionally, math anxiety has been

shown to occupy working memory resources, further reducing available cognitive capacity during mathematical tasks (Ashcraft & Krause, 2007).

Mathematical Challenges in Children with Hearing Loss

Children with hearing loss represent a distinct population whose mathematical difficulties arise from a different, though partially overlapping, set of cognitive and environmental factors. Research consistently shows that students with hearing loss underperform hearing peers in mathematics across educational levels (Blatto-Vallee et al., 2007; Marschark et al., 2015). These difficulties are not attributable to intellectual impairment but rather to a constellation of factors: communication barriers that limit access to verbal instruction, increased cognitive load from listening effort, language delays that impede mathematical word problem comprehension, and reduced incidental learning of numeracy concepts through everyday conversation (Caemmerer et al., 2016).

A critical insight from the literature concerns cognitive load. In hearing children, cognitive resources are divided between processing academic content and other classroom demands. For children with hearing loss, however, an additional cognitive burden arises from the effort required to decode spoken language - through lip-reading, hearing aids, cochlear implants, or sign language interpretation - before mathematical content can even be processed (Audiology Island, 2024). This compound cognitive load reduces the working memory resources available for mathematical reasoning, creating a compounding disadvantage in conventionally delivered instruction.

Language-dependent mathematics presents a particular obstacle. Deaf and hard of hearing children often arrive at school with less developed mathematical language than hearing peers, due to reduced access to the incidental verbal numeracy learning that occurs in everyday family and social contexts (Pagliaro, 2010). Word problems - central to most mathematics curricula - place disproportionate demands on linguistic processing, further becoming a disadvantage to this population (Blatto-Vallee et al., 2007).

The role of visual-spatial processing in this population merits careful consideration. Research has shown that visual-spatial schematic representations - those that capture relational structure rather than surface pictorial features - are a stronger positive predictor of mathematical problem-solving performance for deaf students than for hearing students (Blatto-Vallee et al., 2007). This finding suggests that instructional approaches that develop and leverage visual-spatial schematic thinking may offer particular benefit to children with hearing loss, even as they should not assume a generalized visual learning advantage (Marschark et al., 2013). Three additional cognitive deficit areas warrant explicit attention in the literature review, as they directly inform the theoretical case for Vedic Mathematics as an alternative instructional pathway. First, children with hearing loss frequently exhibit phonological processing deficits that extend beyond the sensory level: because phonological awareness develops through exposure to spoken language, reduced auditory input during critical developmental periods can impair the phonological loop’s capacity to encode and retrieve arithmetic facts, which in typical instruction

are rehearsed verbally (Marcelino et al., 2019). This deficit is particularly relevant to mathematics because standard approaches to multiplication tables, number bonds, and counting sequences rely heavily on phonological rehearsal strategies. Second, children with hearing loss show disproportionate difficulty on sequential verbal working memory tasks specifically, while often performing comparably to hearing peers on visual-spatial memory measures (Marschark et al., 2013). This asymmetry is theoretically significant: it suggests that the deficit is not in working memory capacity per se, but in the verbal-sequential channel through which conventional mathematics instruction is predominantly delivered. Third, research documents elevated rates of executive function challenges in children with hearing loss, including difficulties with planning, inhibitory control, and cognitive flexibility, partially attributable to reduced access to the verbal scaffolding through which hearing children internalize self-regulatory strategies during mathematical problem-solving (Caemmerer et al., 2016). Taken together, these three deficit clusters-phonological processing, sequential verbal working memory, and executive self-regulation-constitute a coherent cognitive profile that conventional auditory-sequential mathematics instruction is poorly designed to accommodate, and that the visual, pattern-based, non-sequential character of Vedic Mathematics is specifically positioned to address.

Vedic Mathematics: Origins, Structure, and Pedagogy

Vedic Mathematics was systematized by Jagadguru Swami Sri Bharati Krishna Tirthaji Maharaja (1884 -1960) and introduced through his 1965 text of the same name. The system comprises sixteen sutras and thirteen sub-sutras, each a concise Sanskrit phrase encoding a computational principle. Examples include Nikhilam Navatascaramam Dasatah (all from nine and the last from ten) and the Vertically and Crosswise sutra - both of which simplify multiplication and division through pattern recognition and mental shortcuts (Williams, 2002). Critically, the system emphasizes mental computation over written algorithms, pattern recognition over procedural memorization, and multiple solution pathways over single prescribed methods - features that distinguish it sharply from conventional Western mathematics instruction.

Existing Evidence for Vedic Mathematics in Educational Settings

Empirical research on Vedic Mathematics remains nascent. Studies conducted primarily in India and the United Kingdom have reported improvements in calculation speed and accuracy among typically developing students following Vedic Mathematics instruction (Glover, 2002; Trivedi & Sharma, 2013). However, these studies are largely limited by small sample sizes, absence of control groups, and lack of standardized outcome measures. No studies have specifically examined the effects of Vedic Mathematics instruction in children with learning disabilities or hearing loss, underscoring the need for the theoretical framework developed here.

Theoretical Foundations

Our framework integrates five established theoretical perspectives from cognitive science and educational psychology. Figure 1 illustrates the Working Memory Model that underpins several of our core propositions.

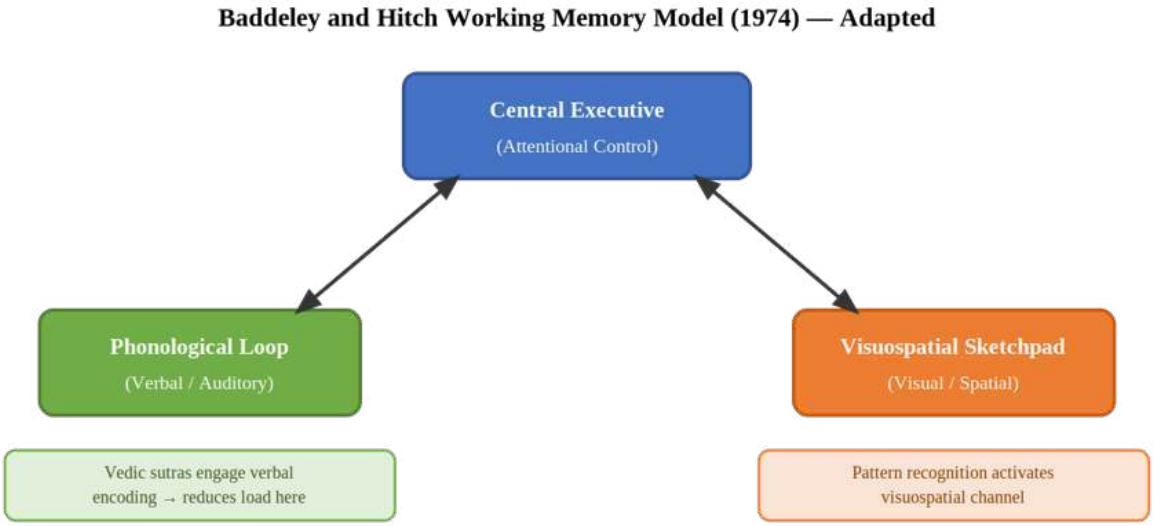


Fig. 1: The Baddeley and Hitch (1974) Working Memory Model, adapted to show how Vedic Mathematics engages both the phonological loop and visuospatial sketchpad - distributing cognitive load across subsystems. This dual-channel engagement is particularly relevant for children with hearing loss, for whom the visuospatial pathway may serve as the primary access route.

Cognitive Load Theory

John Sweller's Cognitive Load Theory (CLT) proposes that working memory has limited capacity, and that instructional designs which minimize extraneous cognitive load enhance learning (Sweller, 1988). For children with learning disabilities - whose working memory capacity is already reduced - conventional multi-step arithmetic algorithms impose high intrinsic and extraneous load simultaneously. For children with hearing loss, this burden is compounded further by the cognitive effort required to access auditory instruction. Vedic Mathematics sutras compress multi-step procedures into single, pattern-based operations, potentially reducing the intrinsic load of arithmetic tasks and freeing working memory resources for deeper mathematical reasoning. This load-reduction effect may be especially pronounced for hearing-impaired students, for whom eliminating the auditory decoding burden through visual instruction represents a direct cognitive intervention.

Working Memory Model

Baddeley and Hitch's (1974) Working Memory Model distinguishes between the phonological loop (verbal and acoustic information), the visuospatial sketchpad (visual and spatial information), and the central executive (attentional control). For children with hearing loss, the phonological loop is functionally compromised as a pathway for mathematical instruction delivered through speech. Vedic Mathematics, by engaging the visuospatial sketchpad

through its pattern-based number recognition and the phonological loop through written or signed encoding of sutras, distributes cognitive processing across subsystems in a way that does not depend on intact auditory processing. This dual-channel engagement aligns with the evidence that visuospatial schematic representation is a stronger predictor of mathematical success for deaf students than for hearing students (Blatto-Vallee et al., 2007).

Information Processing Theory

Information Processing Theory conceptualizes learning as the encoding, storage, and retrieval of information. The pattern-based nature of Vedic Mathematics may facilitate more efficient encoding of arithmetic procedures into long-term memory for all learners - but for children with hearing loss, who have been shown to have difficulties with sequential verbal memory tasks (Marcelino et al., 2019), the non-sequential, holistic pattern recognition that Vedic techniques emphasize may offer a particularly effective encoding route. Repeated practice with Vedic sutras may support automatization of calculation, reducing controlled processing demands over time (Anderson, 1983).

Constructivist Learning Theory

Vygotsky's and Piaget's constructivist frameworks emphasize active knowledge construction through discovery and interaction. For children with learning disabilities hearing loss, who may experience social isolation and reduced participation in conventional classroom discussions of mathematics (Pagliaro, 2010), Vedic Mathematics offers a self-directed, internally driven learning experience that does not depend on verbal group participation. Its multiple solution pathways support independent mathematical discovery, potentially restoring a sense of mathematical agency that classroom exclusion can erode.

Dual Coding Theory

Paivio's (1971) Dual Coding Theory proposes that information encoded through both verbal and visual channels is more robustly represented in memory. For children with learning disabilities and hearing loss, the verbal channel is compromised or absent in its auditory form. However, Vedic Mathematics provides visual encoding through number patterns and spatial arrangements - and for students who use sign language, the embodied, spatial nature of signed arithmetic may provide an additional encoding modality. This multimodal, non-auditory approach to encoding mathematical procedures aligns directly with the compensatory strengths available to children with hearing loss.

A Conceptual Framework: Mapping Populations to Vedic Mathematics Components

Building on the theoretical foundations above, Table 1 presents our conceptual model mapping specific learning and hearing profiles to Vedic Mathematics components, proposed cognitive mechanisms, and expected outcomes. Figure 2 illustrates the sequential cognitive

pathway from instruction to academic outcome. Figure 3 provides a dedicated mapping for children with hearing loss. Figure 4 presents an integrated overview of all five target populations.

Table 1. Mapping of Learning and Hearing Profiles to Vedic Mathematics Components

Population / Difficulty	Vedic Math Component	Proposed Cognitive Mechanism	Expected Outcome
Poor working memory (dyscalculia)	Stepwise mental algorithms	Reduces intrinsic cognitive load	Better calculation accuracy
Slow processing speed	Mental shortcuts and sutras	Faster arithmetic retrieval	Improved math fluency
Low attention / ADHD	Pattern-based learning	Increases engagement and novelty	Better participation
Math anxiety	Simplified computation pathways	Increases self-efficacy	Greater confidence
Phonological deficit (dyslexia)	Visuospatial pattern recognition	Compensatory encoding channel	Improved number representation
Hearing loss — auditory processing deficit	Visual sutras; non-auditory patterns	Bypasses compromised auditory channel	Equitable math access
Hearing loss — increased cognitive load	Compressed, low-load algorithms	Frees working memory from listening effort	Improved reasoning capacity
Hearing loss — language-based problems	Non-symbolic computation	Reduces linguistic processing demand	Better problem-solving
Hearing loss — math anxiety / exclusion	Independent mental computation	Restores agency and self-efficacy	Greater confidence

Fig. 2: Proposed Conceptual Pathway — Vedic Mathematics to Academic Outcome

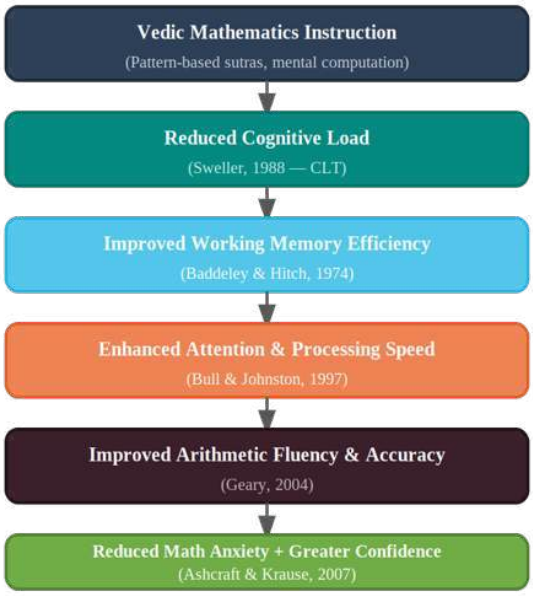


Fig. 2: Proposed sequential cognitive pathway from Vedic Mathematics instruction to sustained academic achievement, grounded in Sweller (1988), Baddeley and Hitch (1974), Bull and Johnston (1997), Geary (2004), and Ashcraft and Krause (2007).

Fig. 3: Hearing Loss, Mathematical Challenges, and Vedic Mathematics as Intervention

Challenge Area	Cognitive Mechanism	Vedic Math Response
Auditory processing deficit	Verbal instruction poorly received	Visual sutras & patterns bypass auditory channel
Increased cognitive load from listening effort	Reduced working memory capacity	Compressed algorithms reduce cognitive demand
Language-based math word problems	Sequential verbal processing difficulty	Non-symbolic / visual pattern-based computation
Math anxiety from classroom exclusion	Diminished self-efficacy & participation	Independent mental math builds confidence

Adapted from: Marschark et al. (2015); Blatto-Vallee et al. (2007); Audiology Island (2024); Ashcraft & Krause (2007)

Fig. 3: Hearing loss challenge areas, associated cognitive mechanisms, and the corresponding Vedic Mathematics response for each. Adapted from Marschark et al. (2015), Blatto-Vallee et al. (2007), Audiology Island (2024), and Ashcraft and Krause (2007).

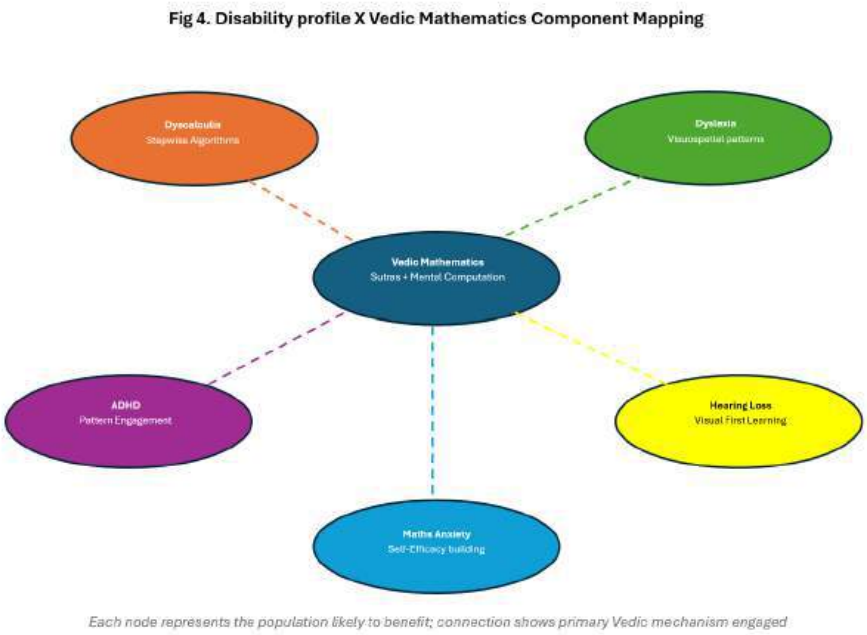


Fig. 4: Integrated disability profile map. Each node represents a population theorized to benefit from Vedic Mathematics; connecting lines indicate the primary Vedic mechanism engaged for each profile.

Vedic Mathematics and Children with Hearing Loss: A Focused Analysis

The case for Vedic Mathematics as an instructional tool for children with hearing loss rests on several converging lines of theoretical argument that merit dedicated elaboration beyond the general framework presented above.

The Auditory Instruction Problem

Conventional mathematics instruction is predominantly auditory and language-dependent: teachers explain procedures verbally, students participate in oral discussions, and word problems require linguistic decoding before mathematical reasoning can begin. For children with hearing loss - whether using hearing aids, cochlear implants, sign language, or oral communication - every layer of auditory instruction represents an additional cognitive demand before the mathematical content itself can be engaged (Audiology Island, 2024). This structural mismatch between the auditory delivery of conventional instruction and the sensory profile of hearing-impaired learners is a primary driver of mathematical underachievement in this population (Caemmerer et al., 2016).

Vedic Mathematics disrupts this dynamic. Its core pedagogical medium is visual and pattern-based: learners observe number structures, identify regularities, and apply sutras that can be represented in written, gestural, or diagrammatic form without dependence on auditory narration. A child with hearing loss engaging with the Nikhilam multiplication sutra, for example, works with visual number relationships that are fully accessible without sound. This

non-auditory accessibility is not incidental to Vedic Mathematics - it is structurally inherent in a system designed for mental, rather than verbally mediated, computation.

Mental Mathematics as Equalizer

Mental mathematics - the capacity to perform calculations internally without written or spoken mediation - represents a domain where children with hearing loss may be placed on a more equitable footing with hearing peers. Unlike classroom discussions or teacher-led verbal explanations, mental calculation is an intrinsically private cognitive act. Vedic Mathematics, with its emphasis on internalized sutras and mental algorithms, trains exactly this capacity. By developing fluency in mental computation through Vedic techniques, children with hearing loss may cultivate a mathematical competence that does not require auditory access to express or demonstrate - reducing the visibility of their disability as an academic barrier.
Furthermore, the research of Blatto-Vallee et al. (2007) demonstrated that when deaf and hard of hearing students do employ visual-spatial schematic representations in mathematical problem-solving - as opposed to surface-level pictorial representations - their performance improves markedly. Vedic Mathematics, by training learners to perceive number patterns as structured, relational schemas rather than arbitrary sequences, may cultivate precisely the schematic visual-spatial thinking that this research identifies as most beneficial for deaf learners.

Sign Language, Embodied Cognition, and Vedic Sutras

For children who use sign language as their primary communication modality, an additional layer of theoretical alignment emerges. Sign languages are spatial, embodied, and visually organized - sharing structural properties with the spatial pattern recognition at the heart of Vedic Mathematics. Research has documented that the manipulation of complex mental images is an area of relative strength for many deaf individuals who use sign language (Emmorey et al., 1993, as cited in Marschark et al., 2013), and that this capacity for mental image manipulation supports geometric and spatial mathematical reasoning. Vedic techniques that engage spatial number arrangements may leverage this embodied cognitive strength, providing an instructional approach that works with, rather than against, the cognitive architecture of signing deaf learners.

Comorbidity of Hearing Loss and Learning Disabilities

A further consideration of clinical relevance is the documented comorbidity between hearing loss and specific learning disabilities. Research indicates that children with hearing loss have elevated rates of co-occurring learning disabilities, including dyscalculia and reading difficulties, potentially arising from shared neurobiological vulnerability or from the downstream effects of language delay on cognitive development (Savegnago et al., 2024). For children carrying this dual diagnosis, the theoretical benefits of Vedic Mathematics proposed for both populations may be additive: the visual-first approach reduces auditory processing demands

while simultaneously reducing the working memory load that dyscalculia imposes on arithmetic tasks. This population may represent the highest-priority subgroup for empirical investigation.

Theoretical Hypotheses

Based on our conceptual framework, we derive the following testable hypotheses for future empirical investigation:

- H1: Vedic Mathematics instruction reduces intrinsic cognitive load during arithmetic tasks across all target populations, as measured by subjective cognitive load ratings and dual-task performance paradigms.
- H2: Pattern-based computation through Vedic sutras enhances long-term memory consolidation of arithmetic procedures, resulting in faster and more accurate fact retrieval.
- H3: Repeated mental calculation using Vedic techniques strengthens executive functioning, particularly planning and inhibitory control, in children with ADHD and dyscalculia.
- H4: Vedic Mathematics instruction significantly reduces cognitive overload in children with hearing loss by bypassing the auditory decoding burden of conventional instruction, resulting in greater working memory availability for mathematical reasoning.
- H5: Children with hearing loss who receive Vedic Mathematics instruction will show greater improvements in arithmetic fluency relative to hearing peers receiving the same instruction, due to the disproportionate cognitive load relief the visual-first approach provides.
- H6: Children with comorbid hearing loss and dyscalculia represent the population most likely to demonstrate the largest absolute gains, as Vedic Mathematics addresses both the auditory access deficit and the working memory burden of dyscalculia simultaneously.

Future Research Directions

This theoretical framework identifies several high-priority directions for future empirical research, organized by study design and target population.

Randomized Controlled Trials

Rigorous randomized controlled trials (RCTs) comparing Vedic Mathematics instruction against conventional mathematics curricula in children diagnosed with dyscalculia, hearing loss, and comorbid profiles represent the most pressing research need. Such trials should employ standardized neuropsychological assessments of working memory, processing speed, and executive function at baseline, and measure arithmetic fluency, accuracy, and math anxiety as primary outcomes. For hearing loss populations, trials should be stratified by communication

modality (hearing aid, cochlear implant, sign language) to identify whether specific subgroups show differential benefit.

Neuroimaging Studies

Functional neuroimaging studies using fMRI or EEG could examine whether Vedic Mathematics instruction produces measurable changes in neural activation patterns associated with arithmetic processing - particularly in the intraparietal sulcus (associated with number sense) and prefrontal cortex (associated with working memory and executive function). For deaf participants, neuroimaging studies may also reveal the degree to which Vedic instruction activates visuospatial processing regions rather than auditory cortices, providing biological evidence for the compensatory pathway proposed here.

Adapted Delivery Formats for Hearing Loss

Research is needed to develop and evaluate Vedic Mathematics instructional materials specifically adapted for children with hearing loss. This includes video-based visual demonstrations of sutras with captioning, signed versions of Vedic algorithms for Deaf signing communities, and tactile or diagrammatic representations of number patterns for children with combined hearing and visual impairments. The development of such materials should proceed in partnership with deaf education specialists and hearing loss advocacy communities to ensure cultural and communicative appropriateness.

Comorbidity Studies

Given the elevated rates of co-occurring learning disabilities in the hearing loss population (Savegnago et al., 2024), dedicated studies examining children with comorbid hearing loss and dyscalculia are warranted. These studies should assess whether the combined theoretical benefits of Vedic Mathematics for both conditions produce additive effects, and should examine the role of hearing loss severity (mild, moderate, severe, profound) as a moderating variable.

Limitations

This paper presents a theoretical framework rather than empirical findings, and several limitations must be acknowledged. The proposed mechanisms are speculative: while grounded in established cognitive theory, they remain untested in populations with learning disabilities or hearing loss. The existing literature on Vedic Mathematics is primarily focused on typically developing learners, and generalization to clinical populations requires empirical validation. With respect to hearing loss, the research base on visual-spatial processing and mathematics achievement is itself contested - the assumption of a generalized visual learning advantage for deaf students has been challenged by several studies (Marschark et al., 2013) - and our framework is careful to ground its claims in schematic rather than general visual processing. Finally, the cultural and linguistic origins of the Vedic system may present implementation challenges that this framework does not fully address.

Conclusion

This paper has developed a theoretical framework for the application of Vedic Mathematics in children with learning disabilities and hearing loss, grounded in five established cognitive and educational theories. We proposed a conceptual model mapping specific population profiles to Vedic Mathematics components and their plausible cognitive mechanisms, derived six testable hypotheses, and identified priority directions for future empirical research. Four original figures provide visual summaries of the Working Memory Model as it applies to this framework, the proposed cognitive pathway from instruction to outcome, the hearing loss challenge-response mapping, and the integrated disability profile map.

The theoretical case for Vedic Mathematics as a supplementary instructional tool for these populations is compelling and merits rigorous empirical investigation. Its emphasis on visual pattern recognition, mental computation, and cognitive efficiency addresses the specific working memory, processing speed, and auditory access deficits documented across dyscalculia, dyslexia, ADHD, and hearing loss. The non-auditory, visually accessible nature of Vedic techniques represents a structural alignment with the needs of hearing-impaired learners that few existing mathematics pedagogies can claim.

For a student researcher whose own journey has been shaped by the intersection of mathematical curiosity and advocacy for children with hearing and learning differences, this framework is not merely academic - it is a call to action for a field that has yet to fully explore the potential of ancient knowledge in service of modern learning equity.

Works Cited

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed., American Psychiatric Publishing, 2013.

Anderson, John R. The Architecture of Cognition. Harvard University Press, 1983.

Ashcraft, Mark H., and Jeremy A. Krause. "Working Memory, Math Performance, and Math Anxiety." Psychonomic Bulletin & Review, vol. 14, no. 2, 2007, pp. 243–248.

Audiology Island. "Understanding the Impact of Hearing Loss on Education and Learning." audiologyisland.com, 16 Nov. 2024, audiologyisland.com/blog/understanding-the-impact-of-hearing-loss-on-education-and-learning/.

Baddeley, Alan D., and Graham Hitch. "Working Memory." The Psychology of Learning and Motivation, edited by Gordon H. Bower, vol. 8, Academic Press, 1974, pp. 47–89.

Blatto-Vallee, Gary, et al. "Visual-Spatial Representation in Mathematical Problem Solving by Deaf and Hearing Students." Journal of Deaf Studies and Deaf Education, vol. 12, no. 4, 2007, pp. 432–448.

Bull, Rebecca, and Rona S. Johnston. "Children's Arithmetical Difficulties: Contributions from Processing Speed, Item Identification, and Short-Term Memory." Journal of Experimental Child Psychology, vol. 65, no. 1, 1997, pp. 1–24.

Butterworth, Brian, et al. "Dyscalculia: From Brain to Education." Science, vol. 332, no. 6033, 2011, pp. 1049–1053.

Caemmerer, Jacqueline M., et al. "Academic Achievement of Children and Adolescents with Hearing Loss." Exceptional Children, vol. 83, no. 2, 2016, pp. 195–212.

Geary, David C. "Mathematics and Learning Disabilities." Journal of Learning Disabilities, vol. 37, no. 1, 2004, pp. 4–15.

Glover, Peter. Teaching the Times Tables - Applying Vedic Mathematics to the School Curriculum. Inspiration Unlimited, 2002.

Marcelino, L., et al. "Cognitive Foundations of Mathematics Learning in Deaf Students." INTED Proceedings, 2019.

Marschark, Marc, et al. "Are Deaf Students Visual Learners?" Learning and Individual Differences, vol. 25, 2013, pp. 156–162.

Marschark, Marc, et al. "Mathematics Achievement of Deaf and Hard of Hearing Students." Journal of Deaf Studies and Deaf Education, vol. 20, no. 2, 2015, pp. 105–123.

Pagliaro, Claudia M. "Mathematics Instruction and Learning of Deaf and Hard-of-Hearing Students." Oxford Handbook of Deaf Studies, Language, and Education, vol. 1, MIT Press, 2010.

Paivio, Allan. Imagery and Verbal Processes. Holt, Rinehart and Winston, 1971.

Rubinsten, Orly, and Rosemary Tannock. "Mathematics Anxiety in Children with Developmental Dyscalculia." Behavioral and Brain Functions, vol. 6, no. 1, 2010, p. 46.

Savegnago, D., et al. "Learning Disabilities in Children with Hearing Loss: A Systematic Review." American Journal of Otolaryngology, vol. 45, no. 5, 2024.

Seligman, Martin E. P. Helplessness: On Depression, Development, and Death. W. H. Freeman, 1975.

Shalev, Ruth S. "Developmental Dyscalculia." Journal of Child Neurology, vol. 19, no. 10, 2004, pp. 765–771.

Sweller, John. "Cognitive Load During Problem Solving: Effects on Learning." Cognitive Science, vol. 12, no. 2, 1988, pp. 257–285.

Tirthaji, Bharati Krishna. Vedic Mathematics. Motilal Banarsidass, 1965.

Toll, Sylke W. M., et al. "Executive Functions as Predictors of Math Learning Disabilities." Journal of Learning Disabilities, vol. 44, no. 6, 2011, pp. 521–532.

Trivedi, K. N., and R. Sharma. "Effectiveness of Vedic Mathematics on the Achievement of School Students." International Journal of Scientific Research, vol. 2, no. 6, 2013, pp. 128–130.

Williams, Kenneth. Discover Vedic Mathematics. Motilal Banarsidass, 2002.

The Volatility Spillover Between Commodity Prices and Stock Market Indices
By Zacharia Mir

Abstract

This study examines volatility spillovers between four commodities (gold, silver, copper, and platinum) and four stock market indices (S&P 500, FTSE 100, Nikkei 225, and MSCI Emerging Markets) using Granger causality tests on daily data from January 2015 to December 2025, split into pre-COVID, COVID, and post-COVID sub-periods. Volatility spillovers strengthened sharply during the COVID-19 period, reducing diversification benefits. After the pandemic, the volatility transmission effect gradually declined and returned to levels similar to those before COVID-19. Gold transmitted no significant volatility to any stock market index during the pandemic, although it received spillovers from the FTSE 100 and the S&P 500, which qualifies its role as a safe haven. These findings provide useful insights for investors, portfolio managers, and policymakers on risk management and diversification strategies during periods of market uncertainty.

Keywords Commodity Price, Stock Market Index, Volatility Spillover, COVID-19, Granger Causality

Introduction

Commodity and equity markets are large enough that shocks in one can move the other. Changes in commodity prices feed into production costs, inflation, and corporate profits, while equity indices reflect price expectations about future growth and economic conditions. Changes in commodity prices affect production costs, inflation, corporate profit, and global economic growth. Understanding how both markets interact is important for investors, portfolio managers, and policymakers because it can improve diversification benefits and risk management strategies. This relationship becomes more important during the period of financial uncertainty, when stronger connections between markets reduce diversification benefits.

This study examines the volatility spillover relationship between four major commodities (gold, silver, copper, and platinum) and four major stock market indices (FTSE, S&P 500, Nikkei 225, and MSCI Emerging Markets) using daily data from 2015 to 2025. To evaluate how market conditions changed over time, the sample is divided into three periods: pre-COVID-19, during COVID-19, and post-COVID-19. Comparing these periods makes it possible to determine whether the pandemic changed the way volatility was transmitted between commodity and stock markets and whether these changes continued after the crisis.

The findings show that volatility spillover became much stronger during the COVID-19 period, indicating greater interconnectedness between commodity and equity markets, meaning diversification opportunities became more limited. However, these effects were temporary. After COVID-19, volatility spillover gradually returned to pre-COVID levels, suggesting the crisis did not permanently change the relationship between commodity prices and stock market indices. This result

also supports the contagion hypothesis, which suggests that global trade crises increase the transmission of volatility across markets.

This paper contributes to existing literature in several ways. First, unlike many previous studies that examine only one commodity or one stock market, this study analyzes four commodities and four stock indices together, providing a broader view of market interconnectedness. Second, by including both precious metals and industrial commodities, this study helps determine the diversification benefits of different types of commodities during the period of market uncertainty. Third, the analysis includes both developed and emerging markets, allowing for a more comprehensive assessment of global volatility spillovers. Finally, by comparing the pre-COVID, COVID, and post-COVID periods, this study shows how crises affect market relationships and provides evidence that investors and policymakers can use when assessing diversification during crises. The remainder of this paper is organized as follows. Section 2 provides an extensive review of the literature. Section 3 describes the data, methodology, and empirical results. Section 4 concludes the paper.

Literature Review

The relationship between commodity price and stock market indices has been extensively examined using a variety of econometric methods. Previous research has applied techniques such as DSTCC-GARCH (Öztek and Öcal, 2017), Granger causality (Hegerty, 2018), DCC-GARCH (Creti et al., 2013; Trichilli et al., 2023; Raina & Bardhan, 2025), ARDL (Chang and Fang, 2022), PCA (Liang et al., 2020), cointegration analysis (Musawa and Mwaanga, 2017), VAR (Mensi et al., 2013; Kang et al., 2024), ADCC-CoVaR (Ghallabi et al., 2025), TVP-VAR (Ramesh et al., 2025; Dhouha Mellouli et al., 2024; Tu and Leatham, 2025), Diebold–Yilmaz spillover framework (Soltani and Abbes, 2025), impulse response analysis (Kocaarslan, 2025), neural network quantile regression (Anwer et al., 2025), and Engle-Granger approaches (Mongale and Eita, 2014). The studies demonstrate that the relationship between commodity and stock markets has been examined from many different perspectives.

Previous studies have also analyzed a wide range of asset classes. These include coffee (Tu and Leatham, 2025), brent oil (Trichilli et al., 2023), platinum (Anwer et al., 2025), currency markets (Kocaarslan, 2025), wheat (Soltani and Abbes, 2025), lumber (Rasmesh et al., 2025), Crude Oil (Raina and Bardhan, 2025), Soybeans (Al-Anezi et al., 2025), Natural Gas (Kang et al., 2024), Copper (Musawa and Mwaanga, 2017), Nickel (Liang et al., 2020), Precious Metal Commodity Index (Chang and Fang, 2022), Sugar (Creti et al., 2013), and Gold (Öztek and Öcal, 2017). Researchers have examined major stock market indices alongside these commodities, including Emerging Stock Market Index (Ghallabi et al., 2025), S&P 500 (Mensi et al., 2013), Johannesburg Stock Exchange (Mongale and Eita, 2014), Nikkei 225 (Dhouha Mellouli et al., 2024), and European Stock Market Index (Hergety, 2018).

Previous studies have examined this relationship across a wide range of countries and regions. Many studies have focused on the United States (Öztek and Öcal, 2017), Europe (Hegerty, 2018), The G20 (Creti et al., 2013), China (Chang and Fang, 2022), Zambia (Musawa and Mwaanga,

2017), South Africa (Mongale and Eita, 2014), India (Raina and Bardhan, 2025), emerging markets (Ghallabi et al., 2025), The G7 (Dhouha Mellouli et al., 2024), and Russia (Soltani and Abbes, 2025).

Early studies suggest that the commodities and stock markets are not always strongly correlated. For example, Öztek and Öcal (2017) find that the relationship becomes much stronger during crisis periods, when higher volatility increases the correlations between asset classes. This suggests commodities provide greater diversification benefits during normal market conditions. Similarly, Creti et al. (2013) show that commodity and stock returns become more closely connected during financial crises, although gold remains an exception by continuing to act as a safe haven asset. Other studies also show that increases in natural resource commodity prices can positively influence stock markets, particularly in emerging markets like China (Chang and Fang, 2022).

The literature also shows that volatility spillovers differ across commodity groups. Precious metals, especially gold, are considered safe haven assets because they protect during periods of financial uncertainty (Raina and Bardhan, 2025; Ghallabi et al., 2025). In contrast, industrial and energy commodities, including WTI Crude Oil, Natural Gas, and copper, tend to have stronger linkages with equity markets (Chang and Fang, 2022; Musawa and Mwaanga, 2017). Al-Anezi et al. (2025) finds agricultural commodities move with financial markets and offer little diversification, while Tu and Leatham (2025) argue they still improve portfolio risk management despite stronger linkages.

Soltani and Abbes (2025) find that investor sentiment plays an important role in explaining spillovers between financial markets and commodities markets, especially during periods of high uncertainty such as the Russia-Ukraine Conflict and the COVID-19 pandemic. However, once investors understood the scope of such conflicts, investor sentiment shifted towards absorbing shocks. For example, resulting from negative investor sentiment originating from fears of the pandemic and global shutdown, overall risk in global commodity markets increased during COVID-19 (Anwer et al., 2025).

Many papers also find strong volatility spillovers between energy commodities and stock markets, especially during crisis periods. For example, during COVID-19, global shutdown concerns rapidly reduced the demand for crude oil as many manufacturing and transportation processes halted worldwide. As a result, crude oil prices crashed during the COVID-19 period along with stock market indices worldwide, signalling a significant volatility spillover between both asset classes (Ramesh et al., 2025; Anwer et al., 2025). Stock markets are the main source of volatility spillovers to commodities, especially energy, with effects rising during crises, while assets like natural gas and agricultural futures remain more stable and serve as better hedges. (Mensi et al., 2013).

This study fills the gap in existing literature by examining and comparing the volatility spillovers between four commodities and stock market indices, including both developed and emerging markets, before, after, and during the COVID-19 period. This paper provides a comprehensive analysis of how supply chain crises and high market volatility influence market interconnectedness and hedging opportunities. The literature review table below presents an extensive review of the relationship between commodity and equity indices.

Table 1. An Extensive Review of Literature on The Relationship Between Stock Markets and Commodity Indices.

Authors	Year	Key Variable	Region	Methodology	Key Findings	Time Period
Öztek & Öcal	2017	S&P-AG, S&P-PM, S&P500	USA	DSTCC-GARCH	Commodity stock links increased during crisis periods.	1990-2012
Hegerty	2018	Stock Market Indices	US & Europe	VAR & Granger Causality	Falling oil prices increase currency pressure, with stock spillovers passing through neighboring markets.	1998-2017
Creti et al.	2013	Coffee, Sugar, Cocoa, Copper, S&P 500	G20	DCC-GARCH	Commodity and stock returns become more connected during financial crises, except for gold	2001-2012
Chang & Fang	2022	Stock and Precious Metal Indices	China	ARDL	Higher commodity prices raise stock market indices in Emerging Markets	2001-2019
Liang et al.	2020	Crude Oil, Nickel, Aluminum	USA	PCA	Grain and soft commodity futures best forecast stock market volatility, especially in recessions	2007-2012
Musawa & Mwaanga	2017	Copper, Oil, LuSE Index	Zambia	ARDL Cointegration	Copper and interest rates have the strongest long-term effects	1991-2019
Mongale & Eita	2014	Gold, Platinum, JSE	South Africa	Engle-Granger	Higher gold and platinum prices are linked to stronger stock market returns	1994-2013
Mensi et al.	2013	S&P 500, Wheat, Brent	USA	VAR-GARCH	Strong volatility spillovers between commodity and stock markets,	2000-2011
Kang et al.	2024	Corn, Soybeans, Natural Gas, S&P 500	USA	VAR	Stock market volatility lowers commodity prices over time	1999-2021
Al-Anezi et al.	2025	Gold, DJI, Soybeans, S&P 500	USA	GARCH-M	Agricultural commodities move with financial markets and offer little diversification, while gold	2015-2024

					helps reduce risk.	
Raina & Bardhan	2025	BSE-500, VIX, Crude Oil	India	DCC-GARCH	Gold provides protection, and oil reflects geopolitical shocks	2000-2023
Ghallabi et al.	2025	Gold, Commodity Index, Emerging Stock Markets	Emerging Countries	ADCC-CoVaR KS test	The relationship between commodities and emerging market equities is unstable and crisis-driven.	2005-2024
Ramesh et al.	2025	S&P 500, Lumber, Corn, Wheat	USA	TVP-VAR	Stock markets transmit volatility to commodities	1990-2022
Mellouli et al.	2024	Nikkei 225, DAX 40, CAC 40, S&P 500, Gold	G7	TVP-VAR	Financial markets show changing and short-term spillovers, with commodities acting as net receivers of shocks.	2019-2024
Soltani & Abbes	2025	MOEX, Nasdaq, Crude Oil, Wheat	USA and Russia	Diebold–Yilmaz	Investor sentiment and asset returns are strongly linked during crises	2017-2023
Kocaarslan	2025	Oil, Gold, Currencies, Equities	USA	DCC, Impulse Response Function	A rising US dollar strengthens uncertainty across major markets and accelerates the spread of risk.	2008-2021
Anwer et al.	2025	VIX, CBOE, Crude Oil	USA	neural network	Risk in global commodity markets increased after COVID-19.	2018-2021
Trichilli et al.	2023	DJI, S&P 500, SSE, Brent Oil	USA and China	DCC-GARCH EVaR	Confirmation bias raises expected losses and market volatility during COVID-19.	2014-2022
Tu & Leatham	2025	Live Cattle, Cotton, Coffee, S&P 500	USA	TVP-VAR	Linkages between agricultural prices and the stock markets remain stronger today.	2000-2024

Data

Daily price data for four commodities and four stock market indices were collected for the period of January 2015 to December 2025. Data for gold (GCW00), silver (SIW00), platinum (PLW00), copper (HGW00), the S&P 500 Index (.INX), FTSE 100 Index (UKX), and Nikkei 225 Index (NI225) were obtained from Google Finance. Data for the MSCI Emerging Market Index (MSCIEF) was extracted from Investing.com. The full dataset was split into three separate periods: the pre-COVID-19 Period (January 2, 2015, to March 13, 2020), the COVID-19 Period (March 16, 2020, to December 31, 2021), and the post-COVID-19 Period (January 03, 2022, to December 31, 2025). These periods were defined based on the World Health Organization's COVID-19 timelines. The volatility for each commodity and stock market index was calculated using the standard deviation of daily log returns over a 30-day rolling window.

Table 2. Descriptive Statistics of Volatility for the pre-COVID-19 Period (January 2, 2015 to March 13, 2020), the COVID-19 Period (March 16, 2020 to December 31, 2021), and the post-COVID-19 Period (January 03, 2022 to December 31, 2025).

Variable	Time Period	Obs	Mean	Std. Dev.	Min	Max
Gold	***Pre***	1208	0.0099	0.0088	0.0043	0.0631
	COVID	455	0.0114	0.0040	0.0054	0.0253
	Post	1003	0.0098	0.0029	0.0044	0.0207
	Total	2666	.0101	.0064	.0043	.0631
Silver	***Pre***	1204	0.0151	0.0071	0.0072	0.0524
	COVID	455	0.0241	0.0091	0.0120	0.0447
	Post	1003	0.0187	0.0044	0.0122	0.0419
	Total	2662	.018	.0074	.0072	.0524
Copper	***Pre***	1204	0.0128	0.0025	0.0077	0.0202
	COVID	455	0.0151	0.0030	0.0091	0.0227
	Post	1003	0.0163	0.0072	0.0084	0.0517
	Total	2662	.0145	.0051	.0077	.0517
Platinum	***Pre***	1202	0.0131	0.0026	0.0081	0.0262
	COVID	455	0.0231	0.0071	0.0138	0.0476
	Post	1003	0.0188	0.0037	0.0131	0.0459
	Total	2660	.017	.0056	.0081	.0476

FTSE	***Pre***	1190	0.0083	0.0032	0.0037	0.0275
	COVID	446	0.0125	0.0077	0.0044	0.0387
	Post	980	0.0075	0.0030	0.0036	0.0166
	Total	2616	.0087	.0046	.0036	.0387
Emerging Markets	***Pre***	1211	0.0087	0.0026	0.0046	0.0203
	COVID	455	0.0112	0.0057	0.0057	0.0321
	Post	1003	0.0095	0.0031	0.0051	0.0188
	Total	2669	.0094	.0036	.0046	.0321
Nikkei	***Pre***	1072	0.0123	0.0048	0.0050	0.0275
	COVID	426	0.0131	0.0056	0.0066	0.0327
	Post	942	0.0128	0.0055	0.0062	0.0349
	Total	2440	.0126	.0052	.005	.0349
S&P 500	***Pre***	1206	0.0080	0.0038	0.0024	0.0351
	COVID	454	0.0130	0.0108	0.0047	0.0519
	Post	1001	0.0104	0.0045	0.0043	0.0258
	Total	2661	.0097	.0061	.0024	.0519

Table 2 presents the descriptive statistics of volatility for commodity and stock market indices for the pre-COVID-19, COVID-19, and post-COVID-19 periods. Overall, average volatility increased during the COVID-19 periods for nearly all commodities and stock market indices, reflecting the higher level of uncertainty within financial markets during the time period. Among the commodities, silver had the largest increase in average volatility during the pandemic, while gold experienced the lowest increase, showing gold is a strong safe-haven asset. For stock market indices, all market indices faced higher levels of volatility during the COVID-19 period compared with both the pre-COVID-19 and post COVID-19 periods. After the COVID-19 period, volatility gradually returned to normal levels when global trade levels drastically improved and the pandemic weakened in strength.

Model

The paper examines the volatility spillover effect between commodity prices and stock market indices using Granger Causality. The paper focuses on four commodity markets: gold, silver, copper, and platinum, and four stock market indices: FTSE 100, MSCI Emerging Markets, Nikkei 225, and S&P 500. Since this paper aims to investigate whether volatility in one market helps to explain the volatility in another market, all variables are converted into volatility

measures before estimating the model. First, the daily logarithmic returns for each commodity and stock market index are calculated as

$R_i = \ln(P_t/P_{t-1})$

Once the logarithmic return is calculated, the volatility is calculated:

$$\text{Volatility}_t = \sqrt{(1/29) \sum_{i=t-29}^{t} (R_i - \overline{R}_t)^2}$$

Where:

- Volatility_t = 30-day rolling volatility at time t
- R_i = Daily log return
- $\overline{R}_t$ = Log return over the previous 30 trading days

This captures total changes in market uncertainty and allows volatility measurement. In this paper, the Granger Causality test is applied to test directional volatility transmission. The model tests whether past volatility helps predict current volatility in other assets. The general model is as follows:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \gamma_1 X_{t-1} + \gamma_2 X_{t-2} + \varepsilon_t$$

Y_t is the volatility of the receiver asset, X_t is the volatility of the source asset, ε_t and is the error return. If the lagged values of X are statistically significant, then volatility in X is said to Granger Cause volatility in Y. In this case, there is evidence of a spillover effect from the source asset to the receiving asset.

H_0: Volatility in the source asset has not Granger Caused in the receiving asset. A p-value below 0.10 means that we reject the null hypothesis, indicating that there is a statistically significant volatility spillover relationship.

The analysis is conducted for total sample figures and for three sample periods: Pre-COVID Period, COVID-19 Period, and Post-COVID Period. This allows the study to measure how spillover changes during periods of financial stress. By comparing these periods, the models provide evidence on whether crises increase market connectedness, reducing diversification benefits between commodities and stock market indices.

Findings

Figure 1. Line Graphs of Volatility for Commodities: Gold, Silver, Platinum, and Copper during three sub-periods: Pre-COVID-19 Period (January 2, 2015 to March 13, 2020), During COVID-19 Period (March 16, 2020 to December 31, 2021), and Post-COVID-19 Period (January 03, 2022 to December 31, 2025).

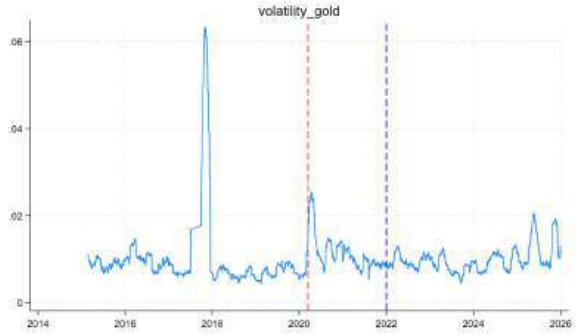
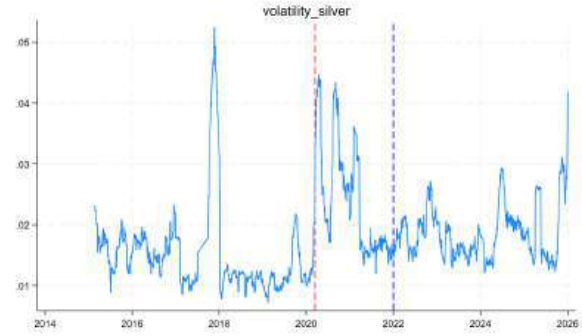
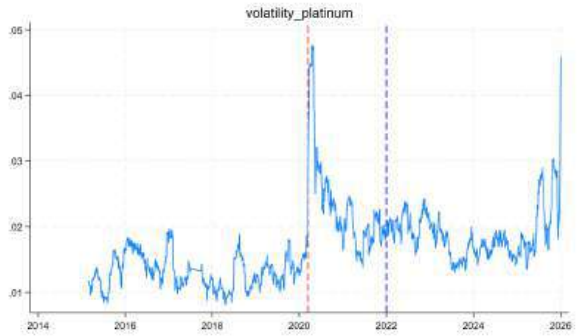
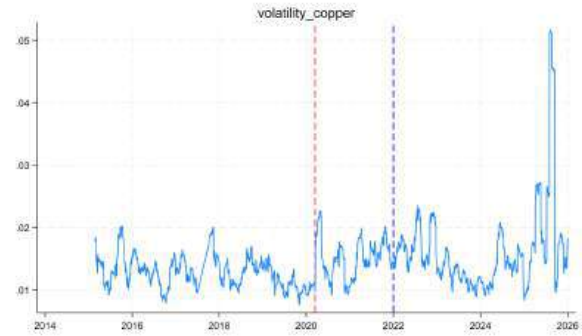
Panel a. Volatility of Gold	Panel b. Volatility of Silver
	
Panel c. Volatility of Platinum	**Panel d. Volatility of Copper**
	

Figure 1 presents the volatility patterns on four major commodities over the full sample period. The vertical red-dashed line represents the beginning of the COVID-19 period. The vertical purple-dashed line represents the end of the COVID-19 period. For all four commodities, the volatility rapidly increased as the COVID-19 period began, peaking near the beginning of the COVID-19 period, with volatility gradually declining as the pandemic progressed until returning to pre-COVID levels once the pandemic ended. However, gold and copper showed a significantly lower initial volatility spike when compared with silver and platinum, consistent with gold acting as a safe haven asset during the period of supply crises and global pandemics.

Figure 2. Line graphs of Volatility for Stock Markets: FTSE 100, MSCI Emerging Markets, Nikkei 225, and S&P 500 during three sub-periods: Pre COVID-19 Period (January 2, 2015 to March 13 2020), During COVID-19 Period (March 16, 2020 to December 31, 2021), and Post COVID-19 Period (January 03, 2022 to December 31, 2025).

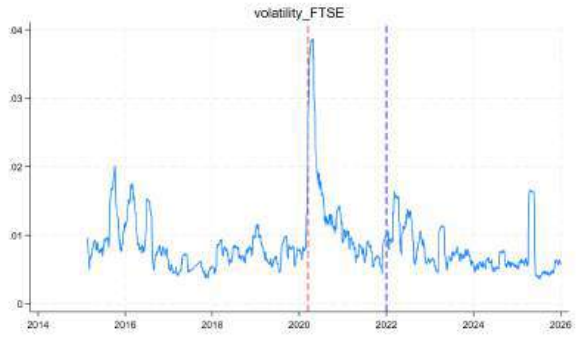
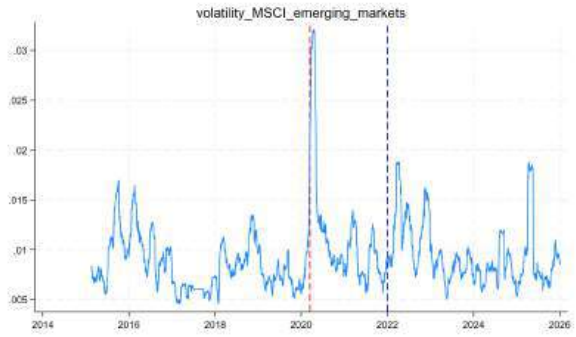
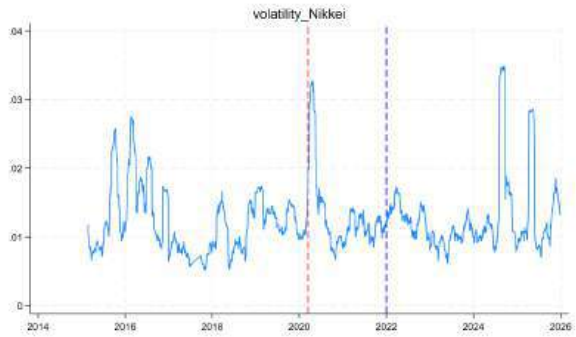
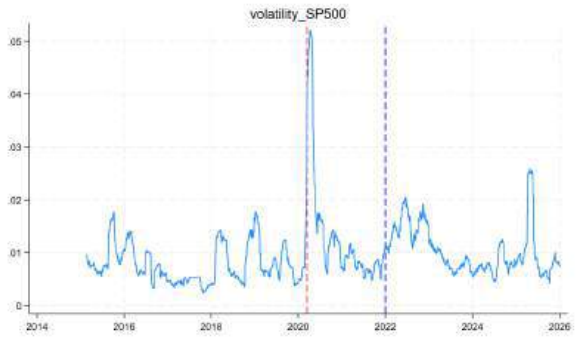
Panel a. Volatility of FTSE	Panel b. Volatility of MSCI Emerging Markets
	
Panel c. Volatility of Nikkei	Panel d. Volatility of S&P 500
	

Figure 2 illustrates the volatility of four stock market indices over the total sample period. The red-dashed line indicates the beginning of the COVID-19 period, while the purple-dashed line indicates the end of the COVID-19 period. All four stock market indices rapidly increased in volatility at the start of the pandemic, reflecting the high level of uncertainty in global financial markets. Volatility peaked during the early stage of the COVID-19 pandemic and gradually declined and returned to similar levels observed before the pandemic period. Among the four indices, the S&P 500 and FTSE 100 showed the highest volatility spikes, while MSCI Emerging Market and Nikkei showed relatively smaller spikes in volatility. Overall, Figure 2 suggests that the COVID-19 pandemic had a significant but temporary effect on both Emerging and developed markets.

Table 2. Volatility Spillover Map Among Commodity and Stock Market Indices Before the COVID-19 Period

	Gold	Silver	Copper	Platinum	FTSE	Emerging Markets	Nikkei	S&P 500
Gold	-	**0.000**	0.461	0.493	0.148	0.837	0.758	0.265
Silver	**0.000**	-	**0.045**	**0.003**	**0.000**	**0.001**	**0.000**	**0.000**
Copper	0.372	0.504	-	0.907	0.661	0.339	0	0.907
Platinum	0.31	0.481	0.425	-	**0.001**	**0.021**	**0.006**	**0.000**
FTSE	0.682	0.841	0.075	0.773	-	**0.000**	**0.001**	**0.000**
Emerging Markets	0.442	0.53	0.581	0.779	0.662	-	**0.006**	**0.000**
Nikkei	0.789	0.92	0.535	0.154	**0.000**	**0.000**	-	**0.000**
S&P 500	0.634	0.324	0.236	**0.006**	**0.000**	**0.000**	**0.007**	-

Table 3. Volatility Spillover Map Among Commodity and Stock Market Indices During the COVID-19 Period

-	Gold	Silver	Copper	Platinum	FTSE	MSCI_emerging_markets	Nikkei	SP500
Gold	-	0.485	0.638	0.647	0.345	0.122	0.402	0.178
Silver	0.646	-	**0.034**	0.477	**0.001**	**0.006**	0.512	**0.093**
Copper	0.569	0.191	-	0.237	0.385	**0.054**	0.372	**0.075**
Platinum	0.891	0.851	0.172	-	**0.000**	**0.000**	0.299	**0.054**
FTSE	**0.013**	**0.044**	**0.095**	**0.000**	-	0.139	0.122	**0.000**
MSCI_emerging_markets	0.527	0.427	**0.035**	**0.006**	0.381	-	**0.003**	**0.000**
Nikkei	0.172	0.197	**0.066**	**0.001**	**0.000**	**0.000**	-	**0.000**
SP500	**0.073**	**0.032**	0.395	0.542	**0.000**	**0.000**	**0.019**	-

Table 4. Volatility Spillover Map Among Commodity and Stock Market Indices After the COVID-19 Period

	Gold	Silver	Copper	Platinum	FTSE	Emerging Markets	Nikkei	S&P 500
Gold	–	**0.007**	0.388	**0.083**	**0.005**	**0.003**	0.251	**0.000**
Silver	**0.000**	–	0.670	0.557	0.355	0.637	0.947	**0.090**
Copper	0.508	**0.030**	–	**0.017**	0.982	0.873	0.450	0.205
Platinum	0.466	0.100	**0.000**	–	0.200	0.260	0.187	0.383
FTSE	0.112	**0.000**	0.576	0.261	–	0.199	**0.023**	**0.000**
Emerging Markets	0.303	**0.033**	0.501	**0.011**	0.125	–	0.801	**0.000**
Nikkei	**0.004**	**0.056**	0.897	0.645	**0.000**	**0.023**	–	**0.000**
S&P 500	0.361	**0.084**	0.265	0.647	**0.036**	0.451	**0.000**	–

Table 5. Volatility Spillover Map Among Commodity and Stock Market Indices, Full Sample Period.

	Gold	Silver	Copper	Platinum	FTSE	Emerging Markets	Nikkei	S&P 500
Gold	-	**0.000**	.572	.151	**.098**	0.159	.600	**.007**
Silver	**0.001**	-	.853	.156	**0.000**	**0.002**	**.001**	**.000**
Copper	0.551	**.012**	-	**.003**	0.789	0.196	.301	**.041**
Platinum	0.322	.101	**.002**	-	**0.000**	**0.002**	.276	**.007**
FTSE	**0.016**	**0.000**	.284	**0.000**	-	**0.000**	**0.000**	**0.000**
Emerging Markets	0.172	**0.010**	0.100	**0.000**	**0.048**	-	**0.011**	**0.000**
Nikkei	**0.089**	.220	.639	**.028**	**.000**	**0.000**	-	**.000**
S&P 500	**0.035**	**.001**	**.000**	**.014**	**.000**	**0.000**	**.000**	-

Tables 2-5 provide the p-values for testing the significance of the spillover relationship between various stock market indices and commodities. Each cell in the matrix represents the p-value of the hypothesis that there is a significant spillover effect from asset 1 (Source) to asset 2 (recipient). The low p values (<0.10) indicate a significant relationship and are bolded and highlighted in red. The high p values suggest the relationship is not statistically significant. The

rows represent where the spillover effect originates. Columns represent the recipients' stock market indices and commodities that are potentially affected by the spillover. The tables reveal multiple significant spillover effects among stock market indexes and commodities, highlighting meaningful interdependencies.

Table 2 presents the volatility spillover relationship during the pre-COVID-19 period (January 2, 2015, to March 13 2020). The findings indicate relatively weaker interconnectedness between commodities and stock market indices compared with later periods. Several significant spillover effects are observed among stock market indices and commodities. Silver is the most consistent transmitter in the pre-COVID period, with significant spillovers toward every stock market index and toward gold, copper, and platinum. Two channels can explain this pattern. First, silver holds a dual role as both a precious and an industrial metal, so its price responds to monetary demand and to manufacturing demand at the same time. A shock to silver volatility, therefore, carries information about both investor risk appetite and industrial conditions, and each channel matters for equities. Second, the investor bases overlap. Traders often treat silver as a risk-on and risk-off signal, and shifts in silver positioning tend to precede portfolio rebalancing across metals and equity indices. Platinum shows a similar but narrower pattern, transmitting significant volatility to all four stock market indices. However, many commodity-to-stock market relationships remain statistically insignificant, suggesting that commodities and equities maintain relatively independent movements during normal market volatility. These findings imply that commodities provide diversification benefits during stable economic periods because volatility transmission across markets was relatively low.

Table 3 presents the volatility relationship during the COVID-19 period. The results indicate stronger interdependencies between commodity and stock markets compared to the pre-COVID period, where market uncertainty increased the transmission of volatility between asset classes. Significant relationships were observed between commodities and stock market indices. For example, Platinum and silver exhibit a significant spillover effect towards FTSE and MSCI Emerging markets. Stock market indices are the dominant transmitters during the pandemic. The four equity rows contain nineteen significant spillovers against nine from the four commodity rows, and the same asymmetry holds in the full sample. This direction matches Mensi et al. (2013), who identify stock markets as the main source of volatility transmission to commodities, and Kang et al. (2025), who find that stock market volatility lowers commodity prices over time. The channel is plausible on two grounds. Equity indices absorb macroeconomic news faster than physical commodity markets because they trade continuously on expectations rather than on spot supply and demand. In addition, the financialization of commodities means that index funds and institutional investors hold commodities inside broader portfolios, so equity losses trigger margin calls and risk reductions that force selling across commodity positions as well. Emerging market indices also received significant spillovers from industrial and precious metals during the pandemic, with copper ($p = 0.054$) and platinum ($p < 0.01$) both transmitting volatility to the MSCI Emerging Markets index. The likely channel runs through index composition. Commodity producers and manufacturers carry heavy weights in emerging market

indices, so uncertainty in metal prices translates directly into uncertainty about the earnings of listed firms. When lockdowns cut industrial production in early 2020, metal demand and the revenues of producer economies fell together. This reading is consistent with Chang and Fang (2022), who find that natural resource commodity prices move Chinese equity indices, and with Musawa and Mwaanga (2017), who show that copper prices have the strongest long-run effect on Zambia's stock exchange. The findings suggest that the COVID-19 pandemic strengthened market connectedness and reduced diversification opportunities across asset classes.

Table 4 presents volatility spillover during the post COVID-19 period. The findings suggest that while the market conditions stabilized after the pandemic, significant volatility transmission remained across commodities and stock market indices. Strong spillover relationships observed from silver towards multiple market indices, including FTSE, MSCI Emerging Markets, and Nikkei. Copper also demonstrates significant interaction between market indices, particularly FTSE and MSCI emerging markets. Additionally, stock market indices continue to display significant interdependencies, especially between the FTSE and the S&P 500. Compared with the COVID-19 period, some spillover effects appeared weaker, implying partial normalization of markets.

Conclusion

The paper examines the volatility spillover among commodities and stock market indices. This paper utilizes daily data for both commodities (gold, silver, copper, platinum) and stock market indices (FTSE 100, MSCI Emerging Markets, Nikkei 225, and S&P 500) from January 2015 to December 2022. The results indicate that interconnectedness between commodity and stock market indices varied. Before the pre-COVID period, volatility spillover was relatively limited. However, during the COVID-19 period, volatility transmission increased among commodities and stock market indices, suggesting that financial stress heightens market connectedness, limiting the efficiency of diversification strategies. The number of significant spillover links in the post-COVID period is close to the pre-COVID count, which suggests the pandemic's effect on market interconnectedness was largely temporary. The composition of those links changed, however: gold, a non-transmitter before and during the pandemic, became a significant transmitter to silver, platinum, and three of the four equity indices afterward. The overall results are in line with the hypothesis that crises transmit volatility between commodities and stock market indices.

This paper has several limitations. The tests are bivariate, so each pair of assets is examined on its own. When a common shock like the COVID-19 pandemic moves many markets at once, a pairwise test can pick up that shared exposure and report it as if one asset were transmitting volatility to another. The spillovers here are best understood as lead-lag relationships in volatility, not direct causal channels. The price data are a third concern, drawn from Google Finance and Investing.com rather than an institutional provider such as Datastream or Bloomberg, and never cross-checked against a second source.

A multivariate model such as DCC-GARCH or TVP-VAR would estimate volatility jointly across all eight assets and replace the rolling standard deviation with a conditional variance, which removes the overlapping-window problem. Rerunning the analysis on institutional data would show whether the results survive cleaner series. The asset set could widen, including agricultural commodities, cryptocurrencies, and currency indices. It could also reach toward the real economy, pairing the commodities and indices with supply and demand measures such as global trade volumes, the trade-weighted dollar, the ISM Manufacturing index, industrial production, and Chinese monthly steel output.

Appendix

Table 1. Data Definitions

Ticker Symbol	Definition	Market Cap	Source
GCW00	Gold Continuous Contract	$27.8T	GOOGLEFINANCE
SIW00	Silver Continuous Contract	$3.8T	GOOGLEFINANCE
MSCIEF	MSCI Emerging Markets Index	$11.7T	Investing.com
.INX	S&P 500 Index	$58.9T	GOOGLEFINANCE
HGW00	Copper Continuous Contract	$8.2T	GOOGLEFINANCE
PLW00	Platinum Continuous Contract	$623.4B	GOOGLEFINANCE
UKX	FTSE 100 Index	$3.5T	GOOGLEFINANCE
NI225	Nikkei 225 Index	$6.05T	GOOGLEFINANCE

Works Cited

Anwer, Z., Khan, A., Naeem, M. A., & Tiwari, A. K. (2025). Modelling systemic risk of energy and non-energy commodity markets during the COVID-19 pandemic. Annals of Operations Research, 345(2), 1193-1227.

Chang, C. L., & Fang, M. (2022). The connectedness between natural resource commodities and stock market indices: Evidence from the Chinese economy. Resources Policy, 78, 102841.

Creti, A., Joëts, M., & Mignon, V. (2013). On the links between stock and commodity markets' volatility. Energy Economics, 37, 16-28.

Ghallabi, F., Ghorbel, A., & Karim, S. (2025). Decoding systemic risks across commodities and emerging market stock markets. Financial Innovation, 11(1), 47.

Hegerty, S. W. (2018). Exchange market pressure, stock prices, and commodity prices east of the Euro. Journal of Economics and Management, (31), 74-94.

Kang, W., de Gracia, F. P., & Ratti, R. A. (2025). Stock market volatility and commodity prices. Macroeconomic Dynamics, 29, e45.

Kocaarslan, B. (2025). Reserve currency and the time-varying link between uncertainties in commodity and financial markets: B. Kocaarslan. Financial Markets and Portfolio Management, 39(3), 415-441.

Liang, C., Ma, F., Li, Z., & Li, Y. (2020). Which types of commodity price information are more useful for predicting US stock market volatility?. Economic Modelling, 93, 642-650.

Mellouli, D., Simo‐Kengne, B. D., Bejaoui, A., & Jeribi, A. (2025). Unveiling the interconnectedness and volatility transmission between real assets and global stock market indices. Review of financial economics, 43(3), 336-382.

Mensi, W., Beljid, M., Boubaker, A., & Managi, S. (2013). Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. Economic modelling, 32, 15-22.

Mongale, I. P., & Eita, J. H. (2014). Commodity prices and stock market performance in South Africa. Corporate Ownership & Control, 11(4), 369-375.

Musawa, N., & Mwaanga, C. (2017). The impact of commodity prices, interest rate, and exchange rate on stock market performance: Evidence from Zambia. Journal of Financial Risk Management, 6(03), 300.

Öztek, M. F., & Öcal, N. (2017). Financial crises and the nature of correlation between commodity and stock markets. International Review of Economics & Finance, 48, 56-68.

Ramesh, S., Low, R. K. Y., & Faff, R. (2025). Modelling time-varying volatility spillovers across crises: Evidence from major commodity futures and the US stock market. Energy Economics, 143, 108225.

Raina, B., & Bardhan, S. (2025). Dynamic Interactions of Geopolitical Risk, Economic Policy Uncertainty and Market Volatility with Stock and Commodity Markets: Evidence from India. International Economic Journal, 39(4), 720-755.

Soltani, H., & Boujelbène Abbes, M. (2026). Unveiling co-movements and spillovers across financial, cryptocurrency, and commodity markets: insights from Google-based investor sentiment. Quality & Quantity, 60(1), 767-790.

Trichilli, Y., Gaadane, S., Boujelbène Abbes, M., & Masmoudi, A. (2025). Behavioural explanations of Expectile VaR forecasting and dynamic hedging strategies for downside risk during the COVID‐19 pandemic: Insights from financial markets. International Journal of Finance & Economics, 30(1), 44-70.

Tu, X., & Leatham, D. (2025). From fields to finance: dynamic connectedness and optimal portfolio strategies among agricultural commodities, oil, and stock markets. International Journal of Financial Studies, 13(3)

Gain-Loss Framing and Loss Aversion in Financial Decision Making
By Pranay Doppalapudi

Abstract

This paper examines how gain-loss framing and loss aversion influence financial decision-making through experimental and field-based research grounded in Prospect Theory (Kahneman & Tversky, 1979). It focuses on the idea that people do not judge outcomes only by their objective value. Instead, they often weigh possible losses more heavily than equal gains, which can shift their risk preferences in many financial decisions.

The literature reviewed includes laboratory experiments, field studies, and applied research in finance, marketing, and education. Across these studies, framing affects behavior in areas ranging from consumer spending to performance under financial incentives. At the same time, the evidence shows that these effects are not fixed. Their strength depends on the situation, the way information is presented, individual differences, and how risk preferences are measured.

The findings support Prospect Theory as a robust framework for understanding deviations from expected utility theory while also highlighting important limitations in the measurement and generalizability of loss aversion. Overall, the evidence suggests that gain-loss framing and loss aversion consistently shape financial decision-making, although their effects depend on both the environment and the presentation of the choice.

Introduction

Traditional economic models assume that people make rational decisions by maximizing expected utility. In other words, this means that people are expected to choose the option with the most benefit. Although this idea has been a foundation of economic theory for decades, behavioral research shows that real financial decisions often do not follow these predictions. Instead of responding only to objective outcomes, people are affected by how those outcomes are presented, sometimes making different choices between options that are mathematically the same.

Developed by Daniel Kahneman and Amos Tversky (1979), Prospect Theory is one of the most influential explanations for these patterns. Over the past several decades, researchers have studied these effects in laboratory experiments, financial markets, consumer behavior, and professional settings. While the research generally shows that framing affects decision-making, questions still remain about how stable these effects are across different contexts, populations, and research methods.

This paper examines how gain-loss framing and loss aversion influence financial decision-making by bringing together evidence from experimental economic studies, behavioral psychology, and field-based research. It also evaluates how consistent these effects are across different settings while considering the limitations of measuring loss aversion. Overall, **gain-loss framing and loss aversion reliably influence financial decision-making by asymmetrically weighting losses over gains, producing predictable shifts in risk preferences across**

experimental studies.

Methodology

This paper uses a structured literature review. It does not run new experiments. Instead, it brings together results from existing studies on gain-loss framing and loss aversion and compares what they show across different settings.

All studies are interpreted through Prospect Theory (Kahneman & Tversky, 1979). This theory is the main framework used to understand why people treat gains and losses differently and why risk preferences change depending on how choices are framed.

The studies were chosen to cover three main types of evidence: controlled lab experiments, field experiments, and studies that focus on measurement and methodology. Only academic research, most of which was peer-reviewed, that directly examined framing effects, risk preferences, or loss aversion in monetary or incentive-based settings was included.

First, laboratory studies were selected to show basic framing effects under controlled conditions. Kirchler, Maciejovsky, and Weber (2004) are included because their study tests how framing affects behavior in experimental asset markets. Experimental asset markets are essentially artificial markets in which participants trade stocks and other assets through a simulated platform. The Asian disease problem studies by Bless, Betsch, and Franzen are also included because they clearly show how people make different choices depending on whether outcomes are framed as gains or losses.

Second, studies conducted in real-world settings were used to determine whether these effects appear outside the laboratory as well. Ganzach and Karsahi (1995) is included because it studies credit card use after gain- and loss-framed messages. Outside of finance, Fryer et al. (2012) are included to show that framing can influence behavior in real incentive-based situations. These studies help show whether framing effects go beyond the lab.

Third, studies on individual differences were included to see how loss aversion varies across people. Schmidt and Traub (2002) and Haigh and List (2005) are used because they compare different groups, including experienced financial professionals, to test whether experience changes loss aversion.

Finally, methodological studies were included to address how risk and loss aversion are measured. Holt and Laury (2002) and Harrison and Rutström (2008) are used because they show that results can change depending on how experiments are designed, especially when it comes to incentives and payoff size.

Together, these sources give a balanced view. They include theory, lab evidence, real world behavior, and critiques of measurement. This makes it possible to compare findings across settings and evaluate how consistent gain-loss framing and loss aversion really are.

Theoretical Background

Traditional economic theories assume that people maximize expected utility when making decisions under uncertainty. Prospect Theory offers a different explanation, arguing that

people evaluate outcomes relative to a reference point rather than in absolute terms (Kahneman & Tversky, 1979). One central part of the theory is loss aversion, which is the tendency for possible losses to feel more important than equal gains. This helps explain why people often make choices that differ from what classical economic models would predict.

Prospect Theory also predicts that framing identical outcomes as gains or losses systematically alters risk preferences. Under gain framing, individuals generally prefer certain outcomes over alternatives that present risk, even when both options have expected value equivalency. Expected value equivalency is when, mathematically, two choices are equivalent when probability is factored in. For example, a guaranteed $50 and a 50% chance of winning $100 each have an expected value of $50, yet most people would prefer the guaranteed outcome. The opposite pattern emerges under loss framing. When choosing between a guaranteed loss and a probabilistic loss with the same expected value, individuals often become more willing to accept risk to avoid a certain negative outcome. Together, these opposing tendencies form the foundation for much of the empirical research on framing effects reviewed in this paper.

Framing

Researchers have closely examined gain framing to see whether people consistently prefer guaranteed outcomes over riskier alternatives when choices are presented positively. Kirchler, Maciejovsky, and Weber (2004) studied whether gain and loss framing could influence investor behavior in an experimental asset market. Participants received either positively or negatively framed information about possible dividend outcomes, even though the information was not actually relevant to the assets' real value. Still, framing affected trading behavior. Participants exposed to positive framing generally held their assets longer and had more optimistic expectations than those who received negatively framed information. The researchers also found evidence of the disposition effect, since participants in gain-framed conditions sold assets more quickly than those in loss-framed conditions. These findings show that investors respond not only to the data itself, but also to how that information is presented, which can influence both trading preferences and the timing of liquidation decisions.

Ganzach and Karsahi (1995) extended the study of framing beyond the laboratory by examining real-life buying behavior among inactive credit card customers. Using Prospect Theory (Kahneman & Tversky, 1979), the researchers tested whether emphasizing losses from non-use would lead to different consumer behavior than emphasizing gains from card usage. Credit card holders who had not used their accounts for three months were randomly assigned to receive either a gain-framed or loss-framed message through a standardized telemarketing call followed by a mailed letter. Consumer behavior was then measured by tracking card use and monetary charges over the next two months. Customers exposed to the loss-framed message used their cards more often and accumulated higher charges than those who received gain-framed messages, with these differences staying consistent throughout the observation period. Loss framed participants also remembered the message more, saw it as more persuasive, and showed stronger involvement with the payment method. Unlike many framing studies done in controlled

laboratory settings, this experiment shows that the way financial information is presented can create measurable changes in real consumer behavior.

Ratcliff et al. (2019) examined how gain and loss framing influenced responses to health messages about physical activity. Gain-framed messages focused on the benefits of exercising, while loss-framed messages emphasized the negative consequences of remaining inactive. Their findings challenged the idea that loss framing is always more effective, showing that the impact of a frame depends on the psychological reactions it creates. For example, loss-framed messages did not consistently produce stronger effects and were influenced by factors such as perceived freedom threat. This suggests that framing does not affect decisions in isolation but instead interacts with how individuals interpret and respond to information. Although this study focused on health behavior rather than financial decisions, it further supports the idea that gain- and loss framing can shape choices by influencing psychological perceptions, helping to explain why investors may respond differently to financial gains and losses.

The loss-framing condition of the Asian disease problem provides one of the clearest demonstrations of the reflection effect described in Prospect Theory. Bless, Betsch, and Franzen (1998) examined how contextual factors influenced responses to the Asian disease problem and found that when outcomes were presented as losses, participants became more willing to choose the probabilistic option over a certain loss. This pattern reflects the steeper slope of the value function in the negative domain, where losses carry greater subjective weight than equivalent gains and encourage risk-seeking behavior as individuals attempt to avoid a guaranteed negative outcome. More importantly, the study shows that this effect is not fixed across contexts. When the same problem was presented as a medical decision, participants displayed the expected increase in risk-seeking behavior. However, framing the scenario as a statistical research problem made the reflection effect much weaker. This suggests that framing influences how people think about a decision, not just whether it is presented as a gain or a loss.

Loss Aversion

Schmidt and Traub (2002) examined the prevalence and strength of loss aversion through a laboratory experiment in which participants repeatedly chose between lotteries involving monetary gains and losses. Instead of relying on hypothetical survey responses, the researchers used these choices to estimate each participant's level of loss aversion. Most participants showed greater sensitivity to losses than to gains. At the same time, the strength of loss aversion varied considerably across individuals, with only some participants consistently showing strong loss averse behavior across every decision task. These findings provide clear behavioral evidence that expected monetary losses influence risky financial decisions, while also showing that loss aversion can vary significantly between people.

Haigh and List (2005) investigated whether extensive market experience reduces loss aversion by comparing professional futures and options traders with undergraduate students in an experimental asset market. Participants repeatedly made investment decisions under conditions designed to measure myopic loss aversion, in which individuals evaluate investments over short

time horizons and place greater weight on potential losses than equivalent gains. Contrary to expectations, professional traders exhibited significantly greater myopic loss aversion than the undergraduate participants despite their extensive experience in financial markets. Rather than eliminating the behavioral patterns predicted by Prospect Theory, repeated exposure to financial decision-making appeared to leave them largely intact. The persistence of loss aversion among experienced traders suggests that these tendencies extend beyond inexperienced or hypothetical decision-makers and remain relevant even in high-stakes financial environments.

The influence of loss aversion also extends beyond financial markets into professional settings. Fryer et al. (2012) examined whether framing financial incentives as potential losses rather than gains would affect teacher performance in a randomized field experiment. Participants were assigned either to a traditional gain treatment, in which bonuses were paid after performance targets were achieved, or to a loss treatment, in which teachers received the bonus at the beginning of the year and were required to return it if those targets were not met. Although both groups were eligible to receive the same financial incentive, teachers in the loss-framed condition produced substantially larger improvements in student mathematics achievement. The authors attributed this difference to loss aversion rather than the value of the incentive itself, arguing that framing the bonus as something that could be lost generated a stronger motivational response. By demonstrating these effects in a large, randomized field experiment, the study provides evidence that loss aversion influences professional decision-making outside controlled laboratory environments.

Patterns and Synthesis

According to the studies reviewed, several clear patterns emerge regarding the influence of gain-loss framing and loss aversion on decision-making. While losses generally exert a stronger influence than equivalent gains, the evidence also shows that the strength, and sometimes even the direction, of these effects depends on contextual factors, framing structure, and individual differences.

One of the most consistent findings is the asymmetry in risk preferences across gain and loss frames. Across the studies reviewed, individuals consistently became more risk-averse when outcomes were framed as gains and more risk-seeking when they were framed as losses, reflecting the value function proposed by Prospect Theory. However, the work of Bless, Betsch, and Franzen demonstrates that these flips in preference are not automatic. Instead, they depend on contextual cues that shape how individuals interpret a decision, suggesting that framing effects arise not simply from whether an outcome is positive or negative, but from the way it is mentally represented.

A second pattern is the persistence of framing effects outside the laboratory. Studies by Ganzach and Karsahi (1995), Kirchler et al. (2004), and Fryer et al. (2012) show that framing can meaningfully influence consumer behavior, investment decisions, and professional performance even when objective incentives remain unchanged. The consistency of these findings across different settings suggests that framing effects are not confined to hypothetical

decision tasks but extend to economically meaningful situations with tangible consequences.

The literature also indicates that loss-framed messages often produce stronger behavioral responses than gain-framed messages, supporting the broader principle of loss aversion. At the same time, this relationship is not universal. Ratcliff et al. (2019) found that gain-framed messages sometimes generated greater reactance and perceived threats to autonomy than loss framed messages, illustrating that emotional and psychological responses can influence the effectiveness of different framing strategies.

Another recurring theme is that people respond to potential losses in different ways. Schmidt and Traub (2002) found major differences in loss aversion across participants, while Haigh and List (2005) showed that even experienced financial professionals can remain strongly loss averse. Together, these findings suggest that loss aversion is a reliable behavioral pattern, but its strength varies across individuals and should not be treated as fixed.

Taken together, the evidence points toward a more nuanced view of Prospect Theory. Gain-loss framing and loss aversion consistently influence decision-making, but their effects come from a mix of context, cognitive evaluation, and individual characteristics rather than acting like universal behavioral laws.

Financial Applications

The findings on gain-loss framing and loss aversion have important implications for financial decision-making, especially in investment behavior, consumer finance, and strategic economic interactions. In financial markets, the way information is presented can influence how investors interpret the same information, leading to predictable differences in trading behavior even when the underlying fundamentals do not change. When outcomes are framed as gains, investors generally become more risk-averse, preferring certainty and avoiding extra uncertainty. When the same outcomes are framed as losses, however, they become more willing to take risks to avoid realizing those losses.

These behavioral tendencies help explain several well-documented market phenomena. One example is the disposition effect, in which investors sell profitable assets too early while holding losing investments for too long, resulting in less efficient portfolio allocation (Kirchler et al., 2004). More broadly, the framing of dividends, returns, or investment performance can influence investor expectations and market demand without changing an asset's objective value. This illustrates that investment decisions are shaped not only by the fundamentals of finance, but also by the way those fundamentals are communicated and perceived.

Loss aversion also has important implications for consumer financial behavior. Ganzach and Karsahi (1995), for example, found that framing credit card usage in terms of losses from non-use produced substantially higher spending than gain-framed messages. Their findings suggest that financial institutions can influence consumer behavior through the presentation of incentives, even when the underlying economic benefits remain identical. Similar patterns appear in other applied settings. In education, loss-framed incentive programs generated stronger improvements in performance than equivalent gain-framed incentives, indicating that the

possibility of losing a reward often motivates behavior more effectively than the opportunity to earn one.

The influence of risk preferences extends beyond individual financial decisions into strategic interactions as well. Murnighan et al. (1988) found that risk aversion can significantly affect bargaining outcomes, as individuals who are more sensitive to uncertainty are often willing to accept less favorable agreements to avoid the possibility of getting nothing. This shows that the behavioral tendencies linked to loss aversion shape not only individual financial choices but also negotiations in which outcomes depend on multiple participants.

Taken together, these findings highlight that financial decision-making is influenced by more than just objective economic information. How choices are framed can meaningfully alter behavior, even when the underlying incentives remain unchanged, making psychological framing an important consideration in both financial theory and practice.

Limitations

Although the literature provides strong evidence for framing effects and loss aversion, several methodological limitations challenge the absolute certainty of these findings. One concern is that measured risk preferences are highly sensitive to elicitation methods and the incentives used. This raises questions about whether the observed behavior reflects stable psychological tendencies or features of the experimental design itself.

Holt and Laury (2002) show that estimated levels of risk aversion vary significantly with payoff size and whether the incentives are real. Participants became much more risk-averse when real financial consequences were introduced, which suggests that preferences measured in laboratory settings are not fixed and can depend on the structure of the experiment. This raises the possibility that some measured loss aversion reflects the task itself rather than deeper cognitive preferences.

Harrison and Rutström (2008) further highlight how estimates of risk preferences can vary depending on experimental design, incentive structure, and measurement approach. As a result, estimates of risk preferences from laboratory studies may vary across different situations, making it more difficult to determine how well experimental findings apply to real-world financial behavior.

Together, these findings show an important methodological challenge in the framing literature. Although experimental studies often show preference reversals under controlled conditions, differences in incentive compatibility, payoff size, and task design can affect how strong the observed effects are. This makes direct comparisons across studies more difficult and limits how confidently laboratory estimates can be applied to real-world financial behavior.

These limitations do not undermine Prospect Theory itself, but they do suggest that its empirical application depends heavily on how risk preferences are measured. Future research could combine the control of laboratory experiments with the realism of real-world studies to better understand whether loss aversion is a consistent part of human decision-making or whether it changes depending on the situation people are in.

Conclusion

Despite the limitations, the evidence reviewed in this paper suggests that gain–loss framing and loss aversion meaningfully shape financial decision-making. Prospect Theory helps explain why people often depart from expected utility theory, especially under uncertainty. Across both laboratory and field studies, losses tend to feel more significant than equivalent gains, which leads to consistent shifts in risk preferences across many decision settings.

At the same time, these effects are not absolute. Framing does not influence everyone in the same way, and its impact depends on context, how information is presented, and how individuals interpret the choices in front of them. Differences across participants, including those seen in studies of professional traders and real-world decision-making, show that loss aversion varies by person and situation.

There are also important limits in how these findings are measured. Laboratory experiments provide strong evidence for framing effects, but results can change depending on incentives, study design, and how decisions are tested. Because of this, it is not always straightforward to apply lab findings directly to real financial behavior.

Overall, the research supports a more nuanced view of Prospect Theory. Gain-loss framing and loss aversion are reliable patterns in decision-making, but they are not fixed rules that apply the same way in every situation. They interact with context, experience, and individual differences. Future research that combines controlled experiments with more realistic, real-world settings will help clarify when these behaviors stay consistent and when they change.

Works Cited

Bless, H., Betsch, T., & Franzen, A. (1998). Framing the framing effect: The impact of context cues on solutions to the "Asian disease" problem. European Journal of Social Psychology, 28(2), 287–291. https://doi.org/10.1002/(SICI)1099-0992(199803/04)28:2<287::AID-EJSP861>3.0.CO;2-U

Fryer, R. G., Levitt, S. D., List, J. A., & Sadoff, S. (2012). Enhancing the efficacy of teacher incentives through loss aversion: A field experiment (Working Paper No. 18237). National Bureau of Economic Research. https://doi.org/10.3386/w18237

Ganzach, Y., & Karsahi, N. (1995). Message framing and buying behavior: A field experiment. Journal of Business Research, 32(1), 11–17. https://doi.org/10.1016/0148-2963(93)00038-3

Haigh, M. S., & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. The Journal of Finance, 60(1), 523–534. https://doi.org/10.1111/j.1540-6261.2005.00737.x

Harrison, G. W., & Rutström, E. E. (2008). Risk aversion in the laboratory. In J. C. Cox & G. W. Harrison (Eds.), Risk aversion in experiments (pp. 41–196). Emerald Group Publishing. https://doi.org/10.1016/S0193-2306(08)00003-3

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. American Economic Review, 92(5), 1644–1655. https://doi.org/10.1257/000282802762024700

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291. https://doi.org/10.2307/1914185

Kirchler, E., Maciejovsky, B., & Weber, M. (2004). Framing effects, selective information and market behavior: An experimental analysis. Journal of Economic Psychology, 25(4), 497–509. https://doi.org/10.1016/S0167-4870(03)00081-6

Murnighan, J. K., Roth, A. E., & Schoumaker, F. (1988). Risk aversion in bargaining: An experimental study. Journal of Risk and Uncertainty, 1(1), 101–124. https://doi.org/10.1007/BF00055566

Ratcliff, C. L., Jensen, J. D., Scherr, C. L., Krakow, M., & Crossley, K. (2019). Loss/gain framing, dose, and reactance: A message experiment. Risk Analysis, 39(12), 2640–2652. https://doi.org/10.1111/risa.13379

Schmidt, U., & Traub, S. (2002). An experimental test of loss aversion. Journal of Risk and Uncertainty, 25(3), 233–249. https://doi.org/10.1023/A:1020923921649

Behavioral and Lifestyle Predictors of Obesity: A Quantitative Analysis of Dietary Habits, Physical Activity, and Technology Use Using Classification Modeling By Valerie Wu

Abstract

Obesity is a serious health problem that affects millions of people worldwide. While genetics play a role, lifestyle choices like diet, exercise, and daily habits are believed to be major factors. This study used a dataset of over 2,100 people to examine which lifestyle behaviors best predict a person's obesity level. Using ANOVA, classification models (Logistic Regression, Decision Tree, Random Forest), and regression analysis, I identified the strongest predictors of obesity. Results show that physical activity was lower in higher obesity categories, which confirms my expectations. However, vegetable consumption was highest in the most obese group, which suggests that eating vegetables alone does not prevent obesity if overall diet quality / calorie intake is poor. A significant interaction effect was also found, indicating that the benefit of physical activity is diminished when combined with a high calorie diet. These findings demonstrate that obesity is complex and that multiple lifestyle factors interact, with implications for public health recommendations and the application of data analytics to health research.

Keywords obesity prediction, lifestyle factors, classification modeling, random forest, behavioral health, data analytics, BMI regression

1. Introduction

1.1 Background and Context

Obesity is a complex, multifactorial health condition defined as abnormal / excessive fat accumulation that presents a risk to health. Globally, obesity rates have nearly tripled since 1975, which makes it a major public health crisis. The World Health Organization identifies obesity as a leading risk factor for chronic conditions including type 2 diabetes, cardiovascular disease, hypertension, and several types of cancers.

While genetic predisposition & metabolic factors are important, lifestyle & behavioral choices like diet, physical activity, sleep, and technology use are widely recognized as modifiable risk factors. According to Vahid et al., "One isolated biomarker, therefore, could not enable optimal risk stratification and prognosis for the individual." Researchers increasingly advocate for a "multimodal" approach that combines multiple biological & behavioral markers to better understand & predict obesity. Understanding which specific behaviors are most strongly associated with obesity can inform targeted public health interventions.

1.2 Significance of the Research

Identifying modifiable behavioral factors can help individuals make informed health decisions and guide public health policy. As Görmez et al. shows, machine learning models can

accurately predict obesity levels from lifestyle data, with implications for developing screening tools & personalized interventions. This study's findings may help healthcare professionals & policymakers design more effective prevention strategies by identifying which lifestyle behaviors matter most for obesity risk.

1.3 Research Questions

Primary Research Question: Among adults in Mexico, Peru, and Colombia, which combination of lifestyle & behavioral factors (eating habits, physical activity, water consumption, meal frequency, technology use, transportation method) best predicts an individual's obesity level, and how accurately can we classify individuals into obesity categories using only these behavioral variables?

Secondary Research Question: Are there statistically significant differences in key behavioral factors (such as vegetable consumption frequency, water intake, physical activity frequency, and technology use) between individuals in different obesity categories?

1.4 Hypotheses

Hypothesis 1 (Primary): Physical activity level (FAF) & frequency of vegetable consumption (FCVC) will be the strongest behavioral predictors of obesity category. However, no single factor will dominate, and overall classification accuracy will be moderate (55-70%) because genetic & metabolic factors also play important roles.

Hypothesis 2: People in high obesity category will have less physical activity, higher consumption of high calorie foods, and higher technology use. Specifically, I expect a gradient where each increasing obesity level is associated with progressively worse behavioral patterns.

1.5 Thesis Statement

This study demonstrates that lifestyle & behavioral factors, particularly physical activity, dietary habits, and technology use, can predict obesity levels with moderate accuracy. The analysis will reveal that these factors don't operate in isolation but interact in complex ways, with the combined effect of multiple behaviors being more significant than any single factor alone. Understanding these interactions can inform more effective public health interventions & show the value of data analytics in health research.

1.6 Scope and Objectives

This study uses a dataset of over 2,100 individuals from Mexico, Peru, and Colombia. The data is a mix of real survey responses (23%) and synthetically generated data (77%) created using the SMOTE technique to balance the dataset. The analysis focuses on 16 behavioral & demographic variables. Objectives include building classification models, comparing group differences using ANOVA, test for interaction effects between physical activity & dietary habits, and comparing results across different outcome measures (BMI as continuous vs obesity categories as binary / ordinal).

2. Literature Review

2.1 Obesity is Multifactorial and Demands Multimodal Approaches

Obesity is not caused by a single factor but rather comes from a complex interaction of genetic, environmental, and behavioral influences. According to Danesh Doost et al, genetic background can be a significant factor in obesity's metabolic phenotypes, but modifiable factors like unhealthy dietary patterns, physical inactivity, improper sleep habits, smoking, and alcohol consumption are strongly linked to an increase risk of metabolic abnormalities. This distinction between nonmodifiable genetic factors & modifiable lifestyle factors is important for understanding prevention & intervention strategies.

Vahid et al emphasizes that obesity is a "complex and multifactorial complication" involving age, sex, diet, physical activity, psychological & emotional status, genetic make up, epigenetics, and gut microbiota. Their comprehensive scoping review shows a crucial insight: "One isolated biomarker, therefore, could not enable optimal risk stratification and prognosis for the individual." Instead, the authors advocate for a "multimodal" approach that combines biomarkers from different domains, including behavior attributes like dietary patterns, physical activity, sleep habit, and smoking status, along with genetic, epigenetic, and metabolic markers.

Aryannezhad et al. further support this multifactorial understanding by showing that changes in both diet quality & physical activity over time are independently associated with changes in adiposity. Their cohort study of over 7,200 UK adults found that improvements in diet quality (measured by Mediterranean diet score) & physical activity (measured by physical activity energy expenditure) were both negatively associated with changes in body fat & visceral adipose tissue. Simultaneous improvements in both were associated with the greatest adiposity benefit.

Görmez et al., also emphasize the multifactorial nature of obesity in their machine learning study, achieving 93.67% accuracy using CatBoost. Their explainable AI analysis found key predictors including age, weight, height, & specific food patterns. This further reinforces that obesity prediction requires considering multiple interacting factors.

2.2 The Power of Data Driven Approaches

The dataset I'm using for this study comes from a publicly available source specifically designed for data mining & machine learning applications. According to Mendoza Palechor & de la Hoz Manotas, the dataset contains 17 attributes & 2,111 records labeled with the class variable NObesity (obesity Level), allowing classification into 7 categories (2019). The authors note that the data "can be used for different tasks in data mining such as: classification, prediction, segmentation & association," which validates my choice of classification modeling techniques (Mendoza Palechor & de a Hoz Manotas 2019).

A key feature of this dataset is the use of SMOTE (Synthetic Minority Over sampling Technique) to balance the classes. The original 485 survey responses were unevenly distributed across obesity categories, which would have biased machine learning models toward the most frequent categories. SMOTE generated synthetic examples to create a balanced dataset of 2111

records, ensuring that the models could learn to distinguish all categories effectively (Mendoza Palechor & de a Hoz Manotas 2019).

Aryannezhad et al. also demonstrate the importance of rigorous measurement approaches, using objective sensors to measure physical activity rather than relying on self reported data. This underscores the value of high quality data in obesity research, though my study relies on the available self reported data.

2.3 Established Relationships Between Lifestyle Factors & Obesity

The relationship between lifestyle behaviors & obesity is well shown in the literature. Physical activity consistently appears as a protective factor. Danesh Doost et al. found that moderate-to-vigorous physical activity was "identified as a protective factor for metabolic health among obese subjects in most of the studies." Aryannezhad et al. provide compelling evidence: each 1-SD increase in physical activity energy expenditure was associated with a decrease in body fat of 1.40 kg and a decrease in visceral adipose tissue of 108 g (2025).

Dietary patterns also play a critical role. Danesh Doost et al. highlight that healthy dietary patterns are closely associated with metabolic health, while unhealthy patterns characterized by energy dense foods & sugary beverages are positively associated with metabolically unhealthy obesity. Aryannezhad et al. found that each 1-SD increase in Mediterranean diet score was associated with a decrease in body fat of 0.47 kg. Most importantly, simultaneous improvements in both diet quality & physical activity were associated with the greatest adiposity benefit.

Görmez et al. further confirm the importance of lifestyle factors, identifying "Weight" & "Height" as the most important features, followed by "Age", "FAF" (physical activity frequency), "TUE" (technology use/time on devices), & dietary factors like "CAEC" (consumption of food between meals). Their analysis found that both physical characteristics & lifestyle habits are needed for accurate obesity prediction.

2.4 Interaction Effects Between Diet & Physical Activity

A key theme emerging from recent research is the interaction between diet & physical activity. Aryannezhad et al. specifically tested for interaction between changes in diet quality & physical activity. While they found no statistical evidence for a synergistic interaction (the combined effect was additive rather than multiplicative), their joint association analysis showed that "simultaneous increases in MDS and PAEE were associated with a greater magnitude of decrease in adiposity." This means that although diet & physical activity may not amplify each other's effects beyond what would be expected from their individual contributions, addressing both behaviors together makes the best outcomes.

Aryannezhad et al. conclude, "Although improvements in either behavior alone were associated with lower adiposity over time, addressing both behaviors simultaneously could be associated with a substantial increase in the magnitude of potential adiposity benefits." This shows that interventions should target both diet & physical activity rather than focus on one to

the exclusion of the other.

2.5 Machine Learning & Explainable AI in Obesity Prediction

Görmez et al. represent a significant advancement in obesity prediction by integrating machine learning with explainable AI techniques. They compared 6 different machine learning models & used both SHAP & LIME to make their models interpretable. This is important because as the authors say, an important limitation of many machine learning models is their "black box" characteristic which makes the decision making process impossible for healthcare personnel & patients to understand (Görmez et al. 2025).

2.6 Gaps in Existing Research and How My Study Fits In

Even though researchers now agree that obesity is caused by many factors, many studies still look at individual behaviors one at a time or use basic correlation analysis. While Danesh Doost et al. & Vahid et al. both recommend using integrative, multimodal approaches to study obesity risk, there is still a gap between what they suggest and what is actually done in practice (2025; 2026).

My research addresses this gap in several ways: (1) I use classification modeling to test the combined predictive power of multiple lifestyle variables; (2) I conduct ANOVA to identify statistically significant differences in specific behaviors across obesity categories; (3) I explicitly test for an interaction effect between physical activity & dietary habits; and (4) I compare results across different outcome measures (7 categories, 2 categories, and continuous BMI).

3. Methodology

3.1 Data Source & Description

Origin: The dataset was sourced from Mendoza Palechor & de la Hoz Manotas, who originally collected 485 survey responses from individuals in Mexico, Peru, and Colombia. To address class imbalance, the authors used SMOTE resampling, generating synthetic data to create a balanced dataset of 2,111 records. (77% synthetic, 23% real)

Composition: Data from 2,111 individuals aged 14 to 61.

Features: 17 variables including demographics (Gender, Age), physical measurements (Height, Weight, BMI), and behavior variables (FAF, FCVC, CH2O, TUE, FAVC, CAEC, NCP, CALC, SMOKE, SCC, MTRANS, family_history_with_overweight).

Target Variable: NObeyesdad with 7 categories: Insufficient_Weight, Normal_Weight, Overweight_Level_I, Overweight_Level_II, Obesity_Type_I, Obesity_Type_II, and Obesity_Type_III.

3.2 Research Design

This is a quantitative, correlational study analyzing an existing dataset. The study is observational, meaning I can't establish causality but can identify associations & build predictive models.

3.3 Data Analysis Techniques

All analysis was performed in Python. Stages included: (1) Data Preprocessing: loading data, encoding categorical variables, & making simplified 2 category target variable; (2) Exploratory Data Analysis: summary statistics & visualizations; (3) Feature Engineering: correlation analysis & VIF testing; (4) Predictive Modeling: Logistic Regression (baseline), Decision Tree (non-parametric), & Random Forest (ensemble). BMI was excluded from features to avoid data leakage.

4. Results

4.1 Descriptive Statistics & EDA

The dataset contained 2111 records with 18 variables after preprocessing. Approximately 23% (485 records) came from real survey responses, while 77% (1626 records) were synthetically generated using SMOTE. The mean age was 24.3 years (range 14-61), mean BMI was 29.7 (overweight range), FAF averaged 1.01 on a 0-3 scale, and (FCVC) averaged 2.42.

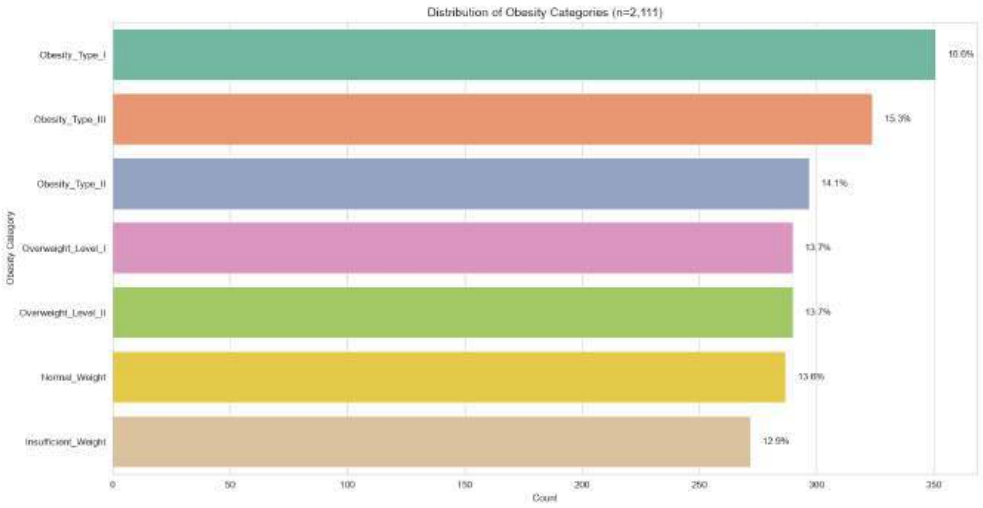


Figure 1: Distribution of Obesity Categories (n=2, 111)

Figure 1 shows the balanced distribution of obesity categories: Obesity_Type_1 (16.6%), Obesity_Type_3 (15.3%), Obesity_Type_2 (14.1%), Overweight_Level_1 & Overweight_Level_2 (13.7% each), Normal_Weight (13.6%), and Insufficient_Weight 12.9%.

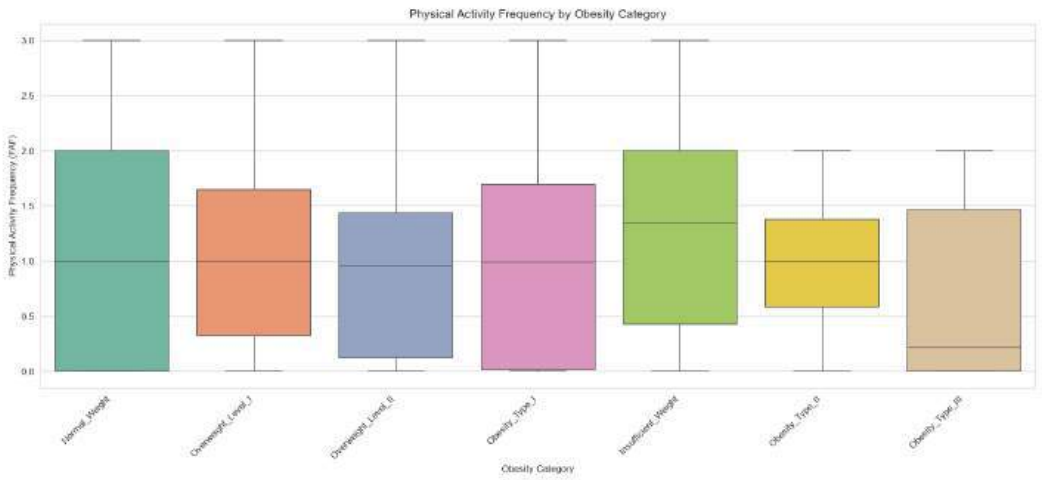


Figure 2: Physical Activity Frequency by Obesity Category

Figure 2 shows a clear negative gradient for physical activity: as obesity level increases, FAF decreases. Obesity_Type_III shows the lowest median FAF (0-0.5), while healthy weight categories show the highest (1.5-2.0).

Figure 3: Correlation Matrix of Key Variables

Figure 3 presents the correlation matrix. The strongest positive correlation with obesity_ordinal was Weight (r=0.913), followed by Age (r=0.283), FAVC_num (r=0.248), and FCVC (r=0.228). The strongest negative correlations were FAF (r=-0.200) and TUE (r=-0.108).

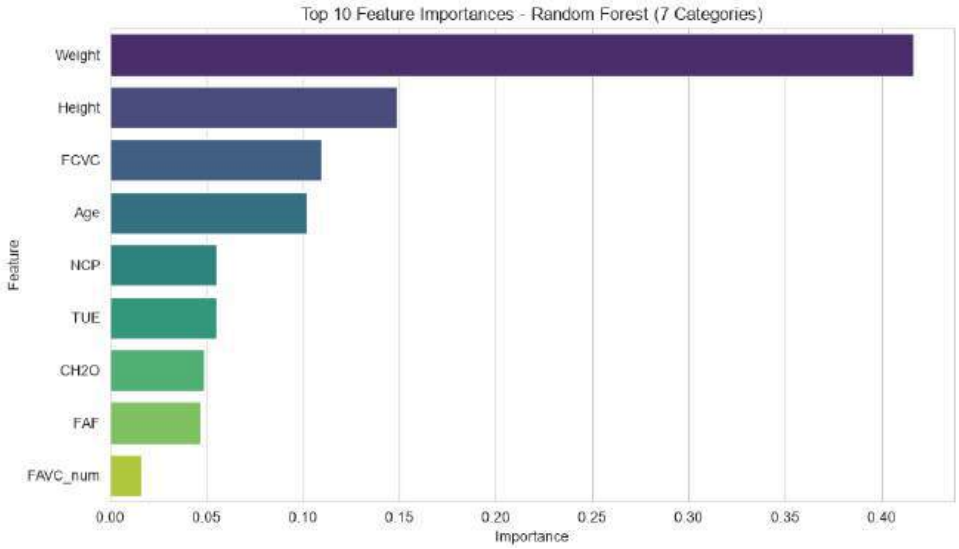


Figure 4: Top 10 Feature Importances - Random Forest (7 Categories)

Figure 4 shows Random Forest feature importance: Weight (41.7%), Height (14.9%), FCVC (10.9%), Age (10.2%), with smaller contributions from NCP, TUE, CH2O, FAF, and FAVC_num. Behavioral variables combined accounted for approximately 30%, while Weight and Height accounted for approximately 57%.

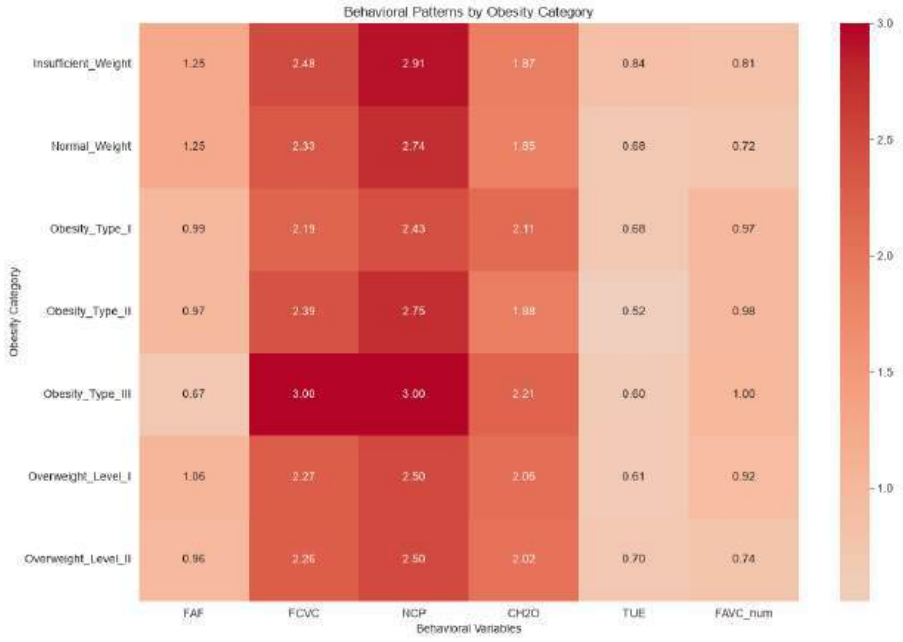


Figure 5: Behavioral Patterns by Obesity Category (Heatmap)

Figure 5 (Behavioral Patterns Heatmap) reveals distinct differences across categories. Insufficient_Weight and Normal_Weight have the highest FAF values (1.250 and 1.247), while Obesity_Type_III has the lowest (0.665). Surprisingly, FCVC is highest in Obesity_Type_III

(3.0). FAVC_num is also highest in Obesity_Type_III (0.997), meaning nearly all individuals in this category consume high-calorie foods frequently.

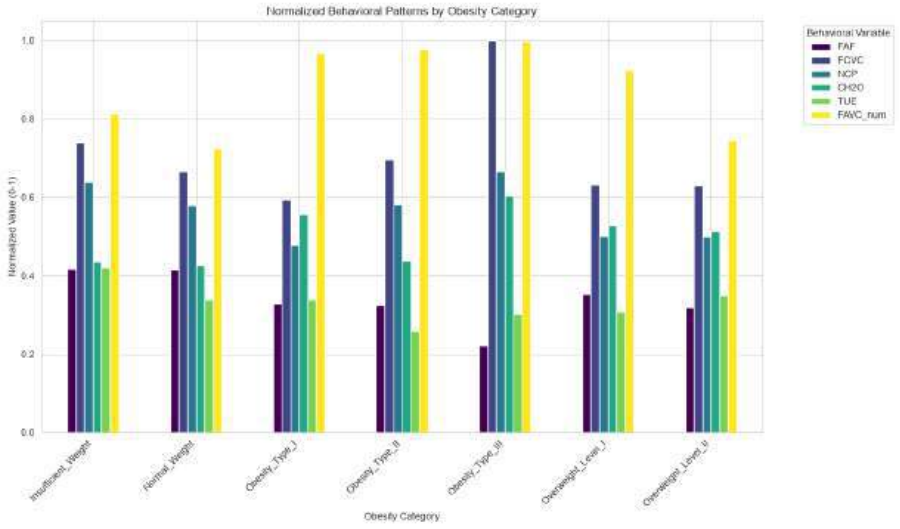


Figure 6: Normalized Behavioral Patterns by Obesity Category

Figure 6 (Normalized Patterns) reinforces that Obesity_Type_III shows the lowest normalized FAF, highest normalized FAVC_num, and highest normalized FCVC.

4.2 Classification Model Performance

Three classification models were trained using features: Age, Height, Weight, FAF, FCVC, NCP, CH2O, TUE, and FAVC_num (BMI excluded to avoid data leakage).

Model Comparison (7 Categories):

Model	Accuracy
Logistic Regression	87.23%
Decision Tree	94.33%
Random Forest	95.98%

The Random Forest model performed consistently across all categories with precision and recall above 0.90. Figure 7 (Confusion Matrix) shows most errors occurred between adjacent obesity categories, as expected for incremental severity levels.

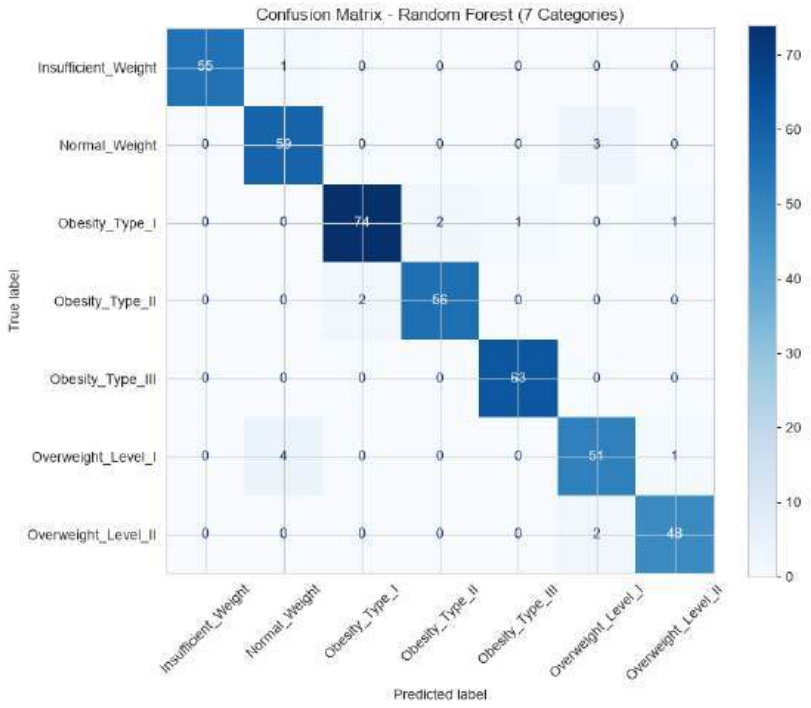


Figure 7: Confusion Matrix - Random Forest (7 categories)

4.3 ANOVA & Post Hoc Test Results

ANOVA tested whether behavioral factors differed across the 7 obesity categories. All variables showed statistically significant differences ($p < 0.001$):

Variable	F-statistic	p-value	Significant
FAF	17.484	7.653e-20	Yes
FCVC	112.315	3.732e-123	Yes
NCP	26.812	6.259e-31	Yes
CH2O	16.171	2.837e-18	Yes
TUEE	7.877	2.069e-08	Yes

FCVC showed the largest differences across categories, followed by NCP, FAF, CH2O, and TUE.

Tukey HSD post hoc tests for FAF revealed significant differences between healthy weight categories and most overweight/obese categories. Notably, Obesity_Type_III had significantly lower physical activity than all other categories. No significant differences were found between Insufficient_Weight & Normal_Weight (p=1.0) or between adjacent obesity

categories (Obesity_Type_I vs. Obesity_Type_II, p=1.0).

4.4 Interaction Model Results

A linear regression model predicted BMI using FAF, FAVC_num, and their interaction term (FAF_FAVC). The model was significant ($F(3,2107)=65.81$, $p<0.001$, $R^2=0.086$).

Coefficient	Estimate	Std Error	T-value	p-value	95% CI Lower	95% CI Upper
Const	24.7626	0.817	30.295	<0.001	23.16	26.366
FAF	-0.3974	0.518	-0.767	0.443	-1.413	0.618
FAVC_num	7.2301	0.862	8.387	<0.001	5.54	8.921
FAF_FAVC	-1.2181	0.56	-2.175	0.03	-2.317	-0.12

The main effect of FAF was not significant (p=0.443), but the interaction term was significant (p=0.030). The negative coefficient indicates that the protective effect of physical activity on BMI is stronger when individuals do NOT consume high-calorie foods frequently.

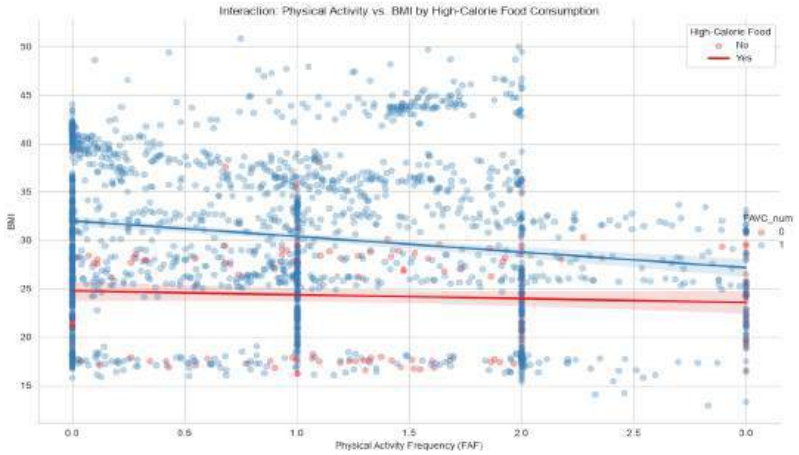


Figure 9: Interaction: Physical Activity vs. BMI by High Calorie Food Consumption

Figure 9 visualizes this interaction, showing a steeper downward slope for the "no high-calorie food" group compared to the "yes" group.

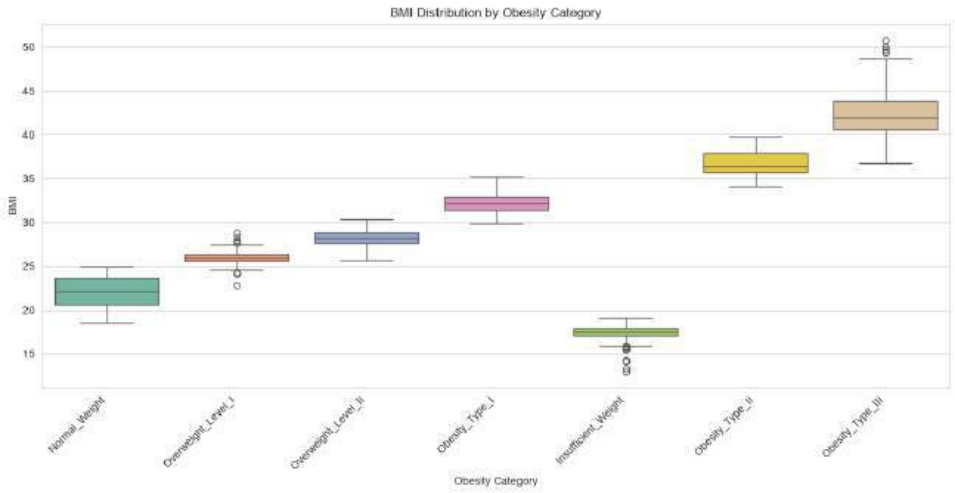


Figure 10: BMI Distribution by Obesity Category

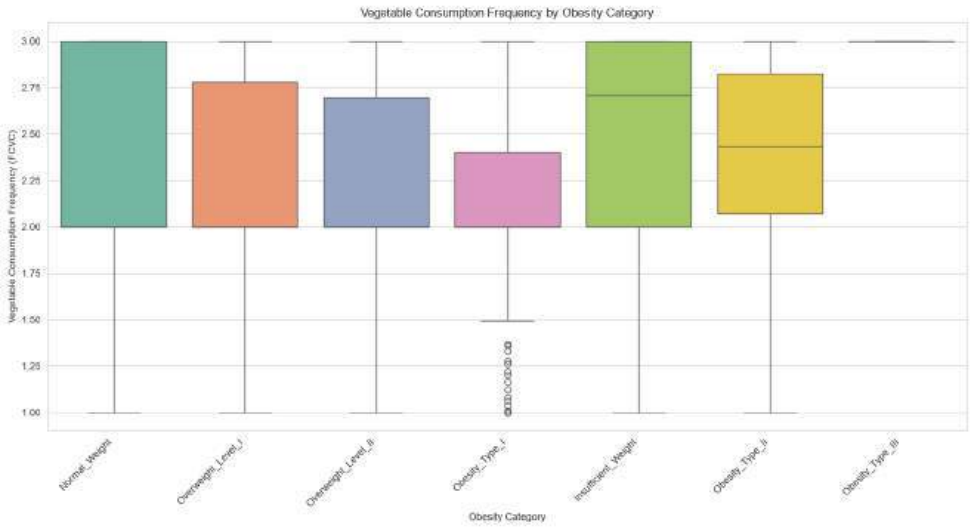


Figure 11: Vegetable Consumption Frequency by Obesity Category

Figures 10 and 11 show BMI and FCVC distributions by obesity category, confirming expected BMI gradients and the unexpected finding that FCVC is highest in Obesity_Type_III.

5. Discussion

5.1 Interpretation of Key Findings

Prediction Accuracy & Primary Hypothesis: The Random Forest model achieved 95.98% accuracy, far exceeding my initial hypothesis of 55-70%. However, Weight dominated feature importance (41.7%), meaning behavioral variables alone would likely have lower predictive power. This aligns with my hypothesis that no single factor dominates and that genetic/metabolic factors also play important roles.

The Role of Physical Activity and the “Lifestyle Gradient”: ANOVA results confirm Hypothesis 2, showing a clear gradient: FAF decreased across increasing obesity categories, with Obesity_Type_III having significantly lower physical activity than all other groups. Tukey HSD tests showed that while physical activity separates healthy weight individuals from those with obesity, it may not distinguish between adjacent obesity categories. This is consistent with Danesh Doost et al.'s finding that moderate to vigorous physical activity is a protective factor for metabolic health. (2025)

Vegetable Consumption and Obesity: The unexpected finding that FCVC was highest in Obesity_Type_III suggests that eating vegetables alone is insufficient to prevent obesity if overall diet quality is poor. Obesity_Type_III had both the highest FCVC and the highest FAVC_num (0.997), meaning nearly all individuals in this category consume high-calorie foods frequently. This may also reflect measurement issues or social desirability bias.

The Interaction Effect: The most important finding is the significant interaction between FAF and FAVC_num on BMI (p=0.030). The negative interaction term indicates that the protective effect of physical activity is stronger for those who do NOT consume high-calorie foods frequently. For those with poor diets, the benefits of exercise are diminished. This is consistent with Aryannezhad et al.'s finding that simultaneous improvements in both diet and physical activity yield the greatest adiposity benefits. (2025)

5.2 Comparison to Existing Research

The findings are broadly consistent with existing research. The protective effect of physical activity aligns with Danesh Doost et al. The Random Forest accuracy (95.98%) compares favorably to Görmez et al.'s 93.67% using CatBoost on the same dataset. Feature importance rankings are similar: body measures dominate, followed by behavioral factors. The interaction effect supports the multimodal approach advocated by Vahid et al. My findings show that this is not just a theoretical recommendation, the combination of diet & physical activity has a greater effect than either factor alone.

5.3 Limitations

Correlation vs. Causation: As an observational study, I can only claim association, not causation.

Self-Report Bias: All behavioral data is self-reported and subject to inaccuracies. As Aryannezhad et al. demonstrated, objective measures provide more accurate measurements. (2025)

Generalizability: The data is specific to Mexico, Peru, and Colombia, with a mean age of 24.3 years, limiting generalizability to other populations and age groups.

Synthetic Data: 77% of the data was generated synthetically using SMOTE. While this helps balance classes and improves model training, synthetic data may not fully capture real-world complexity.

5.4 Suggestions for Future Research

Future research should include a wider range of variables (sleep duration, stress levels, genetic markers, socioeconomic status) and use objective measures of behaviors rather than self report.

6. Conclusion

6.1 Summary of Findings

This study examined relationships between lifestyle factors & obesity using a dataset of 2,111 individuals from Mexico, Peru, and Colombia. The Random Forest model achieved 95.98% accuracy in predicting 7 obesity categories, with Weight (41.7%) and Height (14.9%) as the most important predictors. ANOVA showed all behavioral variables differed significantly across categories, with FCVC showing the strongest differences ($F=112.315$). The most important finding was the significant interaction effect between physical activity and high-calorie food consumption on BMI ($p=0.030$), indicating that the protective effect of exercise is stronger for those with healthier diets.

6.2 Answering the Research Questions

Primary Research Question: Physical activity frequency, vegetable consumption, meal frequency, technology use, water consumption, and high-calorie food consumption together can predict obesity with high accuracy (95.98% for Random Forest). However, Weight and Height dominated predictions, suggesting that behavioral variables alone may not achieve the same accuracy. When simplified to binary outcome (Normal vs. Overweight_Obese), accuracy reached 98.58%.

Secondary Research Question: Yes. All behavioral variables differed significantly across obesity categories, with FCVC showing the largest differences. Tukey HSD tests confirmed a clear lifestyle gradient for physical activity, with Obesity_Type_III having significantly lower physical activity than all other categories.

6.3 Final Thoughts

This study demonstrates that lifestyle and behavioral factors, particularly physical activity, dietary habits, and their interaction, are powerful determinants of obesity. The significant interaction effect suggests that effective prevention and treatment strategies must address both diet and exercise simultaneously. The high accuracy of machine learning models shows that behavioral data can effectively predict obesity status, with implications for developing screening tools. For individuals, the message is clear: lifestyle choices matter, and healthy behaviors work together to reduce obesity risk.

Works Cited

Aryannezhad, Shayan et al. “Concurrent Changes in Diet Quality and Physical Activity and Association With Adiposity in Adults.” JAMA network open vol. 8,11 e2545232. 3 Nov. 2025, doi:10.1001/jamanetworkopen.2025.45232

Danesh Doost, Haniyeh et al. “Association of Genetic, Environmental, and Nutritional Factors With Metabolic Phenotypes of Obesity: A Scoping Review.” Journal of obesity vol. 2025 8472196. 2 Jul. 2025, doi:10.1155/jobe/8472196

"Estimation of Obesity Levels Based On Eating Habits and Physical Condition ." UCI Machine Learning Repository, 2019, https://doi.org/10.24432/C5H31Z.

Fabio Mendoza Palechor, Alexis de la Hoz Manotas, Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico, Data in Brief, Volume 25, 2019, 104344, ISSN 2352-3409, https://doi.org/10.1016/j.dib.2019.104344. (https://www.sciencedirect.com/science/article/pii/S2352340919306985)

Görmez Y, Yagin FH, Yagin B, Aygun Y, Boke H, Badicu G, De Sousa Fernandes MS, Alkhateeb A, Al-Rawi MBA and Aghaei M (2025) Prediction of obesity levels based on physical activity and eating habits with a machine learning model integrated with explainable artificial intelligence. Front. Physiol. 16:1549306. doi: 10.3389/fphys.2025.1549306

Vahid, Farhad et al. “Multimodal (Bio)Markers and Risk of Obesity - A Comprehensive Scoping Review.” Advances in nutrition (Bethesda, Md.) vol. 17,2 (2026): 100579. doi:10.1016/j.advnut.2025.100579

Juvenile Recidivism and Incarceration: A Comparison of U.S. Juvenile Justice Systems
By Ava Miltenberg

Abstract

Across the United States, juvenile justice policy is built on the principle that young people who break the law deserve support, not just punishment, and should be treated differently than adult offenders. However, in practice, that principle is inconsistently applied. While many jurisdictions adopt federal guidance that promotes diversion and rehabilitation, some continue to impose punitive responses, including trying juveniles as adults or placing them in highly restrictive custodial environments that mirror adult incarceration. These contradictions produce drastically different outcomes from state to state, even when the legal frameworks appear similar on paper.

This paper compares two high-custody states, Wyoming and West Virginia, with two of the lowest, New Hampshire and Vermont, to examine what drives these disparities. If sentencing laws appear similar on paper, why are some states more successful at keeping youth out of detention and preventing reoffending, while others continue to rely on confinement?

This paper argues that these differences are shaped less by the policies themselves and more by how state systems interpret, implement, and measure those policies in practice. These distinctions are critical because they influence how recidivism is recorded, how success is evaluated, and whether youth are rehabilitated or pushed deeper into the justice system. By focusing on implementation rather than intention, this analysis reveals the structural factors that determine outcomes across jurisdictions.

Introduction: Why Outcomes Don't Match the Policies

Juvenile justice systems in the United States are not uniform, but they often claim to follow the same foundational principles. Most states model their sentencing and detention practices on recommendations from federal agencies like the Office of Juvenile Justice and Delinquency Prevention (OJJDP). Across jurisdictions, there is widespread agreement on several guiding principles of juvenile justice. Detention is meant to be used only as a last resort, diversion should be prioritized whenever possible, and minor offenses should not lead to lifelong consequences that follow youth into adulthood. These shared standards are intended to promote consistency across state lines.

However, in practice, states implement these principles in dramatically different ways depending on geography, infrastructure, funding, and local judicial norms.[72] This gap between written policy and lived reality is where the issue becomes more complex and where disparities begin to take shape.

For example, a youth arrested for a nonviolent offense in Vermont may be referred to a restorative justice program and avoid court entirely, instead participating in a process that emphasizes accountability, community repair, and personal development. That same youth, if

[72] Council of State Governments. *Juvenile Justice Policy and Practice Review*. 2020.

arrested in West Virginia under similar circumstances, may be denied pretrial diversion and placed in a juvenile facility, entering a system that prioritizes control over rehabilitation. Both states technically endorse similar values, but their courts, probation systems, and detention infrastructures operate differently, and the outcomes reflect those differences in practice, not in policy.

This variation raises the central question of this paper: Why do juvenile recidivism and incarceration rates differ so significantly between states that appear to have similar sentencing policies? What factors beyond policy account for these differences, and how do those factors shape long-term outcomes for youth?

The research and state-level comparisons in this paper suggest that juvenile justice outcomes are shaped less by statutory policy than by the courts, detention systems, and state agencies that interpret and enforce those policies in practice. Differences in infrastructure, geographic access to services, judicial discretion, and local resource availability interact to produce dramatically different incarceration and recidivism rates across states.
Vermont's juvenile custody rate, for example, is technically 0%, a statistic that on the surface suggests an ideal system. In reality, this number reflects the 2020 closure of the state's only juvenile detention facility, Woodside, which left Vermont without any in-state secure placement option for youth. As a result, young people in custody are now sent to adult jails or to out-of-state facilities, most often in New Hampshire. [73] This shift raises critical questions about what "custody" actually means in official reporting and whether those youth are counted under Vermont's data or absorbed into the numbers of the receiving state.

Meanwhile, New Hampshire's custody rate is listed at 10%, but that figure may be partially inflated by Vermont youth housed there, making the data appear more correctional than it might be if adjusted for in-state residents only. These figures complicate direct state-to-state comparisons and raise larger questions about how regional systems intersect and obscure the true scope of youth confinement.

Because of these inconsistencies, this paper adopts a comparative methodology, examining four states—Wyoming, West Virginia, New Hampshire, and Vermont—to analyze how differences in implementation and data reporting shape outcomes. Rather than assuming policies alone determine success, this approach focuses on how systems operate in practice and how structural conditions influence decision-making at every stage.

Beyond policy on paper, outcomes are shaped by decisions made by state legislatures, budget offices, and executive agencies about funding, facility closures, cross-state agreements, and how agencies define and report custody. The same regional pressures that lead one state to close its facilities can shift the burden onto another, and those consequences often remain hidden in statistics that fail to account for cross-jurisdictional dynamics. Ultimately, justice for youth is determined not only by laws, but by who enforces them, how consistently they are applied, and what infrastructure exists to support them.

---

[73] Vermont Public. "Closure of Woodside Juvenile Rehabilitation Center." 2022. + VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.

According to the most recent national data from the OJJDP, there were over 400,000 juvenile delinquency cases processed in U.S. courts in 2022.[74] The vast majority of these cases did not involve serious violent offenses. Still, tens of thousands of youth were detained, sometimes for technical violations or minor offenses that could have been addressed through alternative interventions.

These decisions are shaped not only by law but by what is available locally: whether a diversion program exists, whether a judge trusts it, whether a probation officer has the capacity to supervise alternatives, or whether a youth has stable housing and family support. When this data is disaggregated by state, it becomes clear that the "justice" a youth receives depends heavily on geography. In Wyoming, the youth custody rate per 100,000 residents is more than four times that of Vermont, even though both states report similar youth crime rates.[75]
The implication becomes unavoidable: outcomes are shaped less by what is written into law and more by what actually happens after an arrest.

Literature Review: What the Research Tells Us

Much of the existing research on juvenile incarceration points to two broad conclusions. First, incarceration alone does little to reduce reoffending. Second, outcomes for youth are shaped more by the systems and supports surrounding them than by the punishment itself. Youth who are detained are significantly more likely to be rearrested within two years than their peers who receive community-based alternatives.[76] This pattern reflects not simply the fact of detention, but what incarceration disrupts: education, housing stability, family relationships, and trust in institutions. Rather than addressing underlying issues, incarceration often deepens them and creates new barriers to stability.

This is not just an individual-level problem; it plays out on a national scale, as some states implement their policies more effectively than others depending on training, political will, and access to resources.[77] States like Vermont and New Hampshire have built the infrastructure necessary to enforce diversion consistently, including state-funded restorative justice panels, youth service bureaus, and school-based intervention programs. Vermont, for example, operates a statewide network of community justice centers that handle low-level offenses through restorative conferencing, victim-offender mediation, and youth accountability boards. These programs do not simply divert youth away from court—they actively engage them in structured processes that promote accountability and reintegration. [78]

New Hampshire similarly relies on Juvenile Court Diversion Programs, which often include substance abuse counseling, mental health services, and mentorship requirements tailored to each youth's specific risk factors. These programs function as alternatives not only to

[74] Office of Juvenile Justice and Delinquency Prevention. *Juvenile Court Statistics 2022*. U.S. Department of Justice, 2022.
[75] Ibid.
[76] The Sentencing Project. *The Changing Landscape of Youth Incarceration*. 2023.
[77] The Sentencing Project. *The Changing Landscape of Youth Incarceration*. 2023.
[78] Vermont Public. "Closure of Woodside Juvenile Rehabilitation Center." 2022.

detention but also to system involvement more broadly, preventing youth from entering cycles of court supervision in the first place. In contrast, in places like West Virginia, gaps in funding, transportation, and local service networks make detention the easiest—or sometimes the only—available response, particularly in rural areas where alternative programs remain limited or inaccessible. In these contexts, detention reflects system constraint more than deliberate policy choice.

Understanding this broader structural context is critical to analyzing why some states outperform others despite having similar laws. The difference lies not in what policies allow, but in what systems can realistically support.

Education also plays a critical role. Bouffard, MacKenzie, and Armstrong's (2000) study found that youth without a GED or access to in-custody schooling were significantly more likely to recidivate.[79] This finding was reaffirmed in a 2023 study published in the *American Journal of Criminal Justice*, which emphasized the connection between structured educational programming and reduced recidivism among confined youth.[80] These studies suggest that education is not simply a supplemental service; it functions as a central stabilizing factor that shapes long-term outcomes.

The literature also emphasizes that what happens after detention is just as important as what happens during it. A 2017 report by the Office of Juvenile Justice and Delinquency Prevention (OJJDP), titled *Juvenile Reentry Literature Review*, explains that youth returning from custody without stable housing, education, or adult supervision are far more likely to reoffend.[81] States that coordinate reentry planning from the moment of intake, rather than waiting until release, consistently report better outcomes.

Vermont and New Hampshire actively implement this model through state-funded coordination teams and partnerships with local agencies. In Vermont, reentry planning involves multi-agency collaboration that includes schools, mental health providers, and probation officers, with services like tutoring, family counseling, and transportation support embedded into the transition process. New Hampshire's Department of Health and Human Services works alongside local juvenile courts and community nonprofits to ensure youth are connected with housing assistance, behavioral health services, and vocational training before and after release.
By contrast, West Virginia and Wyoming often lack these consistent structures. In many cases, youth leave custody with little to no follow-up or coordinated plan, which significantly undermines their ability to reintegrate and avoid system re-entry. This gap between release and support creates conditions where reoffending becomes more likely, not because of individual failure, but because of systemic absence.

---

[79] Bouffard, Jeffrey A., Doris Layton MacKenzie, and Gaylene Armstrong.
"Evaluating the Effectiveness of a Correctional Education Program." *Journal of Correctional Education* (2000).
[80] Springer."Education and Recidivism in Juvenile Justice Systems." *American Journal of Criminal Justice*, 2023.

[81] Office of Juvenile Justice and Delinquency Prevention. *Juvenile Court Statistics 2022*. U.S. Department of Justice, 2022.

Finally, the OJJDP's 2022 *Juvenile Court Statistics* offer a national benchmark for offense types, demographic disparities, and court processing trends. States like Vermont and New Hampshire process far fewer youth through formal court channels,[82] relying instead on structured diversion and community-based alternatives. Meanwhile, Wyoming and West Virginia process youth at higher rates and detain more of them, even for similar offenses.[83] This reinforces the argument that policy may appear consistent across states, but the systems behind that policy produce very different realities.[84]

Interpretation, Labels, and the Impact of Judicial Discretion

One of the most powerful, yet often overlooked, forces shaping youth justice outcomes is how the system defines "violence." Across states, the label "violent offender" triggers automatic detention, longer sentencing eligibility, and exclusion from diversion—even when the actual conduct involves little or no physical harm. According to the Justice Policy Institute, many states define violent crimes so broadly that youth can be classified as violent offenders for threats, verbal altercations, or "non-injurious physical contact".[85] Once applied, this label follows youth throughout the system and significantly increases the likelihood of both detention and recidivism.

The *Labeling Violence* report reinforces this point by documenting how racial bias, prosecutorial discretion, and vague legal definitions produce harsher classifications for Black and Brown youth across multiple jurisdictions.[86] Wyoming and West Virginia both show high levels of racial disproportionality in youth detention.[87] Youth of color in these states are more likely to be detained pretrial and more likely to be labeled as violent, even when their offense profiles mirror those of white peers. Vermont and New Hampshire, by contrast, have implemented stricter thresholds for charging youth with violent offenses and place greater emphasis on discretion at intake, reducing automatic exclusion from diversion pathways.

This issue connects directly to the broader role of judicial discretion. On paper, most states allow judges to divert youth into community-based programs. In practice, however, discretion is shaped by perception as much as by policy. As Danielle Hood's 2019 analysis suggests, judicial decisions often reflect perceived risk rather than the severity of the charge itself.[88] In rural states like Wyoming, where judges encounter fewer successful diversion models

---

[82] Office of Juvenile Justice and Delinquency Prevention. *Juvenile Court Statistics 2022*. U.S. Department of Justice, 2022.
[83] Ibid.
[84] Lefas, Melissa. *Rehabilitating Juvenile Violent Extremist Offenders in Detention: Advancing a Juvenile Justice Approach*. Georgetown University, 2016. + Gagnon, Joseph C. "Characteristics of and Services Provided to Youth in Secure Care Facilities." 2010. + The Sentencing Project. *The Changing Landscape of Youth Incarceration*. 2023. + VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.

[85] Justice Policy Institute. *The Costs of Confinement: Why Good Juvenile Justice Policies Make Good Fiscal Sense*. 2016.
[86] *U.S. Criminal Justice Data*.
[87] *Juvenile Reentry Literature Review*. U.S. Department of Justice.
[88] Hood, Danielle. "Judicial Discretion and Juvenile Justice Outcomes." 2019.

and have fewer resources available, detention often becomes the more predictable or administratively manageable option.

This pattern reveals a deeper structural issue. The system does not apply discretion neutrally; it filters decisions through local conditions, resource limitations, and individual perceptions. The result is not simply variation, but inconsistency that affects which youth receive opportunities for diversion and which do not.

Another key factor lies in how youth are categorized—who is viewed as "redeemable" and who is not. Prosecutorial and probation discretion can either open or close pathways depending on the values and pressures within a given jurisdiction. In states like New Hampshire, probation staff often take on an active rehabilitative role, identifying alternatives and advocating for diversion. In West Virginia, where probation offices are frequently understaffed and overextended,[89] decisions often default to incarceration due to time constraints, caseload pressure, and lack of available programs.

This reinforces a central theme: youth justice outcomes depend less on policy than on people and systems. The same statute, applied in two different contexts, can produce entirely different outcomes. Youth in under-resourced or highly surveilled communities face a higher likelihood of detention, not necessarily because of their actions, but because of the systems surrounding them.

Vermont's post-Woodside system provides a particularly revealing case study. With no in-state juvenile facility, responsibility shifted to courts, probation systems, and community programs to coordinate supervision and alternatives. This shift forced a more decentralized and deliberate approach. While the system has faced challenges, including the controversial use of adult facilities in some cases,[90] It has also reduced reliance on incarceration and prompted investment in community-based responses.

Infrastructure Gaps: Education, Mental Health, and Reentry

Beyond the courtroom, a major driver of outcome disparity is access to core infrastructure, particularly education and mental health care. Gagnon's 2010 national study, *Characteristics of and Services Provided to Youth in Secure Care Facilities*, found that educational programming in juvenile facilities varies dramatically across states[91]. Youth in well-resourced systems receive Individualized Education Plans (IEPs), small class sizes, and structured transition planning. Youth in under-resourced systems may receive minimal instruction, often limited to worksheets or inconsistent classroom time [92](Gagnon, 2010). These differences have direct consequences. A study by Gehring (2000) found that youth who earned a GED or re-enrolled in school after release were significantly less likely to be

[89] Gagnon, Joseph C. "Characteristics of and Services Provided to Youth in Secure Care Facilities." 2010.

[90] Vermont Public."Closure of Woodside Juvenile Rehabilitation Center." 2022. + VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.

[91] Gagnon, Joseph C. "Characteristics of and Services Provided to Youth in Secure Care Facilities." 2010.

[92] Ibid.

rearrested[93]. Education does not simply occupy time; it creates continuity, builds skills, and provides a pathway forward that reduces reliance on the system.

Vermont and New Hampshire have made deliberate efforts to integrate education into both diversion and reentry processes. Vermont's youth services system includes partnerships with local school districts and mental health agencies that extend into aftercare. New Hampshire uses a centralized structure to coordinate education, health services, and juvenile court referrals, allowing for more consistent service delivery.

West Virginia and Wyoming, by contrast, rely more heavily on decentralized, county-level systems with fewer safeguards and less coordination. This fragmentation limits post-release stability and increases the likelihood that youth will cycle back into custody. [94]Mental health access reflects a similar divide. In rural areas of West Virginia, families may travel hours to access care, and some counties lack even a single licensed child psychologist[95]. For youth with trauma histories, behavioral challenges, or substance use needs, this absence of support significantly increases vulnerability to reoffending. Lefas (2016)[96] emphasizes that unmet mental health needs directly correlate with higher recidivism, particularly in systems that respond to behavioral issues with confinement rather than treatment.

Comparative: Four States, Four Systems

To understand how these differences play out in practice, it is necessary to examine the four selected states more closely. Wyoming and West Virginia stand out for their relatively high juvenile incarceration rates compared to Vermont and New Hampshire.

In Wyoming, youth are frequently detained for technical violations such as probation noncompliance or curfew infractions, often at rates above the national average[97]. The state's custody rate remains among the highest in the country[98] despite its relatively small youth population ("U.S. Criminal Justice Data"). This pattern suggests that incarceration functions as a default response, reflecting limited alternatives rather than elevated levels of youth crime. West Virginia's system reflects similar structural challenges. Economic instability, rural isolation, and limited access to behavioral health services shape how the system operates. Many youth in the system come from communities that lack transportation, school-based support, or adequate probation staffing. In this context, detention becomes the most available option—not the most effective one.

---

[93] Gehring, Thomas."Recidivism and Educational Achievement in Juvenile Offenders." *Journal of Correctional Education* (2000).

[94] *Juvenile Reentry Literature Review*. U.S. Department of Justice.

[95] Lefas, Melissa.*Rehabilitating Juvenile Violent Extremist Offenders in Detention: Advancing a Juvenile Justice Approach*. Georgetown University, 2016.

[96] Lefas, Melissa. *Rehabilitating Juvenile Violent Extremist Offenders in Detention: Advancing a Juvenile Justice Approach*. Georgetown University, 2016.

[97] The Sentencing Project. *The Changing Landscape of Youth Incarceration*. 2023.

[98] *U.S. Criminal Justice Data*.

The OJJDP's 2017 findings identify the absence of local services as a major driver of recidivism[99]. Without coordinated reentry planning, accessible treatment, or supervision alternatives, youth released from custody in states like West Virginia face significant barriers to reintegration.

In contrast, Vermont and New Hampshire have built systems that emphasize coordination, prevention, and localized response. Vermont's statewide restorative justice framework diverts a significant portion of youth away from formal court involvement. [100]However, the closure of Woodside complicates this picture. While the state reports a 0% custody rate, youth continue to experience confinement—often outside state lines or, in some cases, in adult facilities.

New Hampshire has invested in behavioral health partnerships and centralized case coordination. Its reported 10% custody rate reflects both in-state youth and some out-of-state placements. At the same time, the system emphasizes early diversion screening and structured intervention. When custody is used, it typically connects to longer-term treatment planning, contributing to lower recidivism and shorter detention durations[101] (Juvenile Court Statistics 2022).

Still, these systems are not without flaws. Vermont's reliance on out-of-state placements and occasional use of adult facilities raises serious concerns about safety and oversight[102] (Vermont Public, 2022; VTDigger, 2022). These outcomes highlight a critical point: reducing custody rates on paper does not eliminate the need for appropriate, youth-specific infrastructure.
Policy Isn't Enough, It's What Comes After

This comparative analysis highlights the limits of evaluating juvenile justice systems based solely on written policy. While Wyoming and West Virginia may have diversion statutes similar to those in Vermont and New Hampshire, those policies are only as effective as the systems that support them.

A 2023 study in the *American Journal of Criminal Justice* reinforces this point, finding that access to education and mental health services during custody significantly predicts post-release success[103] (Springer, 2023). However, these services are not distributed evenly. In under-resourced systems, youth may be placed in facilities that lack certified teachers, on-site therapists, or consistent programming.

By contrast, Vermont and New Hampshire have made sustained efforts to integrate individualized education and trauma-informed care into both custodial and diversion settings.

---

[99] Office of Juvenile Justice and Delinquency Prevention. *Juvenile Court Statistics 2022*. U.S. Department of Justice, 2022.
[100] VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.

[101] Office of Juvenile Justice and Delinquency Prevention. *Juvenile Court Statistics 2022*. U.S. Department of Justice, 2022.
[102] Vermont Public."Closure of Woodside Juvenile Rehabilitation Center." 2022. + VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.
[103] Springer. "Education and Recidivism in Juvenile Justice Systems." *American Journal of Criminal Justice*, 2023.

These differences matter. When youth leave a system with academic progress, mental health support, and community connections, they are far more likely to reintegrate successfully. Educational access remains especially critical. Bouffard, MacKenzie, and Armstrong (2000) found that youth without access to structured learning were significantly more likely to return to the system.[104] Vermont and New Hampshire have incorporated education into diversion and aftercare programs,[105] while Wyoming and West Virginia offer fewer consistent opportunities, particularly in rural areas.

Geography also plays a defining role. In states like Wyoming, youth may live hours away from the nearest program or service provider. Judges and probation officers often cite this lack of access when defaulting to detention[106]. These "logistical incarcerations" do not reflect intent; they reflect structural limitations.

This variation reinforces the central argument: policy matters, but implementation determines outcomes.

Conclusion

Juvenile justice in the United States does not fail because of a lack of policy; it fails because of a lack of alignment between principle and practice. This paper examined why states with similar laws produce different outcomes and found that the answer lies in how those laws are carried out.

Wyoming and West Virginia demonstrate how limited infrastructure, geographic barriers, and resource constraints can undermine reform. Vermont and New Hampshire show that systems built around diversion, coordination, and reentry support can reduce reliance on incarceration and produce more stable outcomes.

At the same time, these differences reveal a deeper structural issue. A young person's experience in the justice system depends not only on their actions, but on where they live, who processes their case, and what resources exist in that jurisdiction. The same offense can lead to diversion in one state and detention in another.

These inconsistencies raise important questions about fairness, measurement, and accountability. As seen in Vermont, a state can report a 0% custody rate while still relying on out-of-state placements or adult facilities. Without consistent definitions, data can obscure reality rather than clarify it.

The research makes one point clear: incarceration alone does not address the underlying factors that lead youth into the system. Education, mental health support, stable housing, and consistent guidance shape whether a young person reoffends.

---

[104] Bouffard, Jeffrey A., Doris Layton MacKenzie, and Gaylene Armstrong. "Evaluating the Effectiveness of a Correctional Education Program." *Journal of Correctional Education* (2000).
[105] Vermont Public."Closure of Woodside Juvenile Rehabilitation Center." 2022. + VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.
[106] Williamson, Emily E. "Probation and Juvenile Courts." *The Annals of the American Academy of Political and Social Science* 20 (1902): 259–67. http://www.jstor.org/stable/1009976.

Reducing recidivism requires more than lowering detention rates. It requires building systems that support youth at every stage—before, during, and after contact with the justice system. That means investing in education, expanding mental health access, strengthening coordination across agencies, and rethinking how discretion operates in practice.

A juvenile justice system should be judged not by how efficiently it processes cases, but by how effectively it prevents future harm and supports long-term stability. Until implementation aligns with intention, the gap between principle and practice will continue to define the system—and young people will continue to bear the consequences.

Works Cited

Bouffard, Jeffrey A., Doris Layton MacKenzie, and Gaylene Armstrong. "Evaluating the Effectiveness of a Correctional Education Program." Journal of Correctional Education (2000).

Gagnon, Joseph C., and Brian Barber. "Characteristics of and Services Provided to Youth in Secure Care Facilities." Behavioral Disorders, vol. 36, no. 1, 2010, pp. 7–19. JSTOR, http://www.jstor.org/stable/43153826.

Gehring, Thomas. "Recidivism and Educational Achievement in Juvenile Offenders." Journal of Correctional Education (2000). https://www.csusb.edu/sites/default/files/recidivismasameasure.pdf

Hood, Danielle. "Judicial Discretion and Juvenile Justice Outcomes." 2019.

Justice Policy Institute. The Costs of Confinement: Why Good Juvenile Justice Policies Make Good Fiscal Sense. 2016. https://justicepolicy.org/wp-content/uploads/2022/02/09_05_rep_costsofconfinement_jj_pS.pdf

Lefas, Melissa. Rehabilitating Juvenile Violent Extremist Offenders in Detention: Advancing a Juvenile Justice Approach. Georgetown University, 2016.

Monahan, Kathryn, Laurence Steinberg, and Alex R. Piquero. "Juvenile Justice Policy and Practice: A Developmental Perspective." Crime and Justice 44, no. 1 (2015): 577–619. https://doi.org/10.1086/681553.

Office of Juvenile Justice and Delinquency Prevention. Juvenile Court Statistics 2022. U.S. Department of Justice, 2022.

Juvenile Reentry Literature Review. U.S. Department of Justice.

Springer. "Education and Recidivism in Juvenile Justice Systems." American Journal of Criminal Justice, 2023.

The Sentencing Project. The Changing Landscape of Youth Incarceration. 2023.

U.S. Criminal Justice Data.

Vermont Public. "Closure of Woodside Juvenile Rehabilitation Center." 2022.

VTDigger. "Out-of-State Placement of Vermont Youth After Woodside Closure." 2022.

Williamson, Emily E. "Probation and Juvenile Courts." The Annals of the American Academy of Political and Social Science 20 (1902): 259–67. http://www.jstor.org/stable/1009976.

Why Do Young Consumers Spend Thousands of Dollars Traveling for Concerts?
By Zexin Liu

There are many large-scale festivals, events, concerts, etc. that charge extravagant prices and yet are extremely well attended, especially by members of Generation Z and Millennials. Several factors are responsible for the willingness of younger consumers to spend large percentages of their income on these events, such as FOMO (fear of missing out), a trend toward spending more on experiences and less on physical goods and services than earlier generations and the increased ease of obtaining financing for expensive events by PayPal and credit card companies such as American Express, Capital One and Chase. This represents a significant contrast with previous generations, such as Generation X and the Baby Boomers, who were less willing to spend on concerts and other live events and were also less willing to spend large amounts of their incomes on leisure activities. Earlier generations also had less access to financing for live events and did not obtain information about events from social media since it did not yet exist.

The continuing advancement of technology, especially in the area of social media, along with increasing opportunities for financing the purchase of relatively expensive live events are making it progressively easier for younger consumers to spend more on these events than earlier generations. In addition, changing preferences for live events relative to physical goods and services is also contributing to this trend. Many financial institutions are basing their marketing strategies on this trend and are attempting to capture a larger share of the market from Gen Z and Millennial consumers.

Introduction

There are several extremely expensive music festivals that take place throughout the year in the United States and many other countries. These festivals can end up costing the attendees several thousand dollars after considering several expenses: the cost of the pass required to enter the event, transportation, housing, food and other incidental expenses. Most of the attendees are very young; for example, according to Vogue Magazine, the average age of attendees at the annual Coachella music festival is between 25 and 28 with 73% between 18 and 34. This puts them in either the Millennial Generation (born between 1981 and 1996) or Generation Z (born between 1997 and 2010.) (A complete list of the generally accepted names of all generations born since the late 19th Century is given in the Appendix.). Changes in the complexity and scope of live events, a trend toward more coverage of these events by social media and improved financing opportunities for these events have contributed to the rapidly growing trend of younger generations spending progressively more on live events compared with older generations. This trend is likely to continue with younger generations (Generation Alpha born between 2011 and 2024 and Generation Beta born from 2025) as social media continues to increase its influence while advancing financial technology improves financing opportunities for expensive live events.

Coachella

One of the best examples of an expensive but well attended live event is Coachella. Coachella is an annual music festival held at the Empire Polo Club in Indio, California. The festival is held over two weekends and includes musicians with a wide variety of styles. The event is open to people of all ages. Attendees may take advantage of on-site camping to avoid the extremely high cost of staying in local hotels or other accommodations, but the cost of attending is still quite high.

In 2026 ticket prices started at $549 - $649 for General Admission Passes with VIP Passes available for $1,199 - $1,399. With the overwhelming demand for tickets, resale prices reached much higher levels, ranging up to $7,000 or more. In addition, housing and food were extremely expensive; the entire experience could cost attendees several thousand dollars each. Anyone who needed to travel substantial distances to attend Coachella potentially spent hundreds or thousands of dollars for airfare. The total cost of attending Coachella can be so great that it is possible to pay for it with an installment plan. Despite the large cost, attendance was estimated to be about 125,000 in 2026.

In addition to Coachella, there are several other music festivals around the world that can be even more expensive. These include the Ultra Music Festival in Miami, Florida; Electric Daisy Carnival in Las Vegas, Nevada and Stagecoach (also held at the Empire Polo Club). Internationally, other famous music festivals include Glastonbury in the U.K., Summer Sonic in Japan and Reading and Leeds in the U.K. By far the most expensive music festival is Secret Solstice in Iceland, which offers VIP passes for $1 million.

The Bureau of Labor Statistics (BLS) reports that the cost of admission to sporting events has risen just under 125% since 2000 while the cost of admission to movies, theaters and concerts has risen by about 110% during the same period.[107]

The high cost of music festivals can be attributed to many factors. These include remote or unique locations, optional VIP packages and the extremely high demand for these events. Most importantly, these festivals can often attract the most popular artists for several reasons. One is the huge payouts that they can afford to give these artists; for example, at Coachella some of the headlining acts can earn up to $5 million per appearance. Further, appearing at such a well-known event provides huge exposure to bands which gives them another reason to perform. In addition, Coachella provides immersive art installations and other attractions, such as pop-up lounges where people can charge their phones and laptops and relax. There are also opportunities to experience games or demos of new high-tech products. Attendees can also sample new products, engage in augmented reality activities and may also meet famous people. These certainly enhance the quality of the experience for attendees. The success of Coachella has also been attributed to detailed, careful planning by the organizers. They have been able to

---

107 https://www.bls.gov/opub/ted/2026/big-games-big-prices-admissions-for-sporting-events-up-123-percent-since-2000.htm

fully utilize the resources available to them at the Empire Polo Club, manage huge crowds and maintain important services such as lighting and sanitation.

Coachella is also on the cutting edge of new trends such as "hybrid performances" which mix live performances with interactive experiences. They provide helpful apps for navigation and make it possible to pay for everything without using cash.

Why are Young People Willing to Spend so Much on Concerts and Festivals?

Concerts are a unique experience that offers many benefits to the attendees. They provide a rare opportunity to hear favorite music performed live, which is a very different experience from listening to recorded music. The combination of live music with high quality sound and lighting provides a memorable experience which can't be achieved elsewhere. In addition, experiencing the music with a crowd of other fans of the same music provides a rarely shared experience. Generation Z and Millennials have reported that concerts provide a sense of community which increases the satisfaction from attending concerts.

The demand for live entertainment has rebounded strongly since the COVID pandemic, with sales now exceeding pre-pandemic levels. Along with the rebound in sales has come a surge in prices, with a new term being coined to describe this trend: "funflation." This refers to the rapidly rising demand for live entertainment driving up prices. Research[108] has shown that the amount that people are willing to pay is based somewhat on the generation that they belong to. Generation Z and Millennials have shown that they are the most willing to spend on live events. People who have attended a concert within the last six months spent an average of $150 for the ticket. For sporting events, the average was $132. These high costs have started to lead consumers to consider alternatives and may indicate that prices for live events may have peaked out.

Research has also shown that younger generations (Generation Z and Millennials) report that being part of a community of fans is an extremely important way to make friends and stay in touch with people compared with much smaller numbers for older generations, such as Generation X and Baby Boomers. The social benefits of attending live events is a major factor in the willingness of Generation Z and Millennials to pay high prices for them.

Generational Attitudes About Spending on Concerts and Other Live Events

The willingness of consumers to sacrifice other goods to attend expensive concerts depends to a large degree on the generation that they belong to. Members of Generation Z and Millennials show a greater willingness to make sacrifices to afford concerts and other live events. Surveys have shown that the younger consumers may be willing to reduce their spending on groceries, dating apps, social media apps, streaming subscriptions, eating out, spending on clothes and even nights out with friends to afford concerts. The same research shows that 18%

[108] "'Funflation' – and Live Event Hype – Goes up Against More Cost-Conscious Consumers", Deloitte Center for Technology Media and Communications, October 8, 2024.

of Millennials are willing to spend $500 on a single concert ticket.[109] This trend has been accelerated by the willingness of PayPal and others to finance these types of purchases, offering installment plans as well as interest-free credit. This follows the general trends toward increased spending on live entertainment in the U.S. economy. Research has shown that in 2024 Americans financed $10 trillion worth of purchases with major credit cards, of which nearly 20% was used to finance travel and entertainment expenses. A disproportionate amount of this spending was due to Gen Z and Millennial consumers.

Due to the rising costs of providing live concerts and other events, alternative financing arrangements have become more widespread in recent years. Research by Lending Tree has shown that 23% of consumers have used buy now, pay later loans to finance concert tickets or festivals. The most frequent users of these financing arrangements are Generation Z and Millennials, with 37% of Generation Z and 35% of Millennials having used these arrangements. This is significantly greater than previous generations; for example, 19% of Generation X (born 1965-1980) and only 3% of Baby Boomers (born 1946 – 1964) used the same financing arrangements.

The Rising Cost of Tickets

The rising cost of tickets in recent years can be attributed to several factors, such as "dynamic pricing", scalping and "bots" (robots.) Dynamic pricing refers to the practice of continually updating prices in real time to reflect changing supply and demand conditions. This is currently practiced by ticket vendors, hotels, airlines, cruise lines, etc. to maximize their revenues. The main advantage to consumers of dynamic pricing is that it may be easier to obtain tickets that would have been sold out with a fixed pricing system. The drawback is that prices can rise very quickly and unexpectedly. Scalping refers to the practice of buying tickets at a relatively low price and then charging a much higher price in the secondary market as the demand for a ticket increases. Bots are programs that are set up to buy large quantities of tickets as soon as they go on sale and then resell them later at a much higher price. They are, in effect, automated scalpers. Although bots are illegal in the United States, the practice is widespread, further pushing prices up for consumers.

Recent research has shown that 37% of Gen Z and 39% of Millennials have spent between $500 and $5000 for a ticket to a live concert or sporting event. Young consumers are also more willing to spend on travel expenses to attend live events. A survey showed that Gen Z consumers are willing to spend up to half of their disposable income on travel. Other research has shown that in the last year 65% of Gen Z and 58% of Millennials have traveled more than 50 miles to attend a live event. These generations are more willing than previous generations to go into debt or else earn extra money to finance their traveling and concert expenses. Overall,

109 "Would You Give up Dining Out or Delete Social Media to Afford Concert Tickets? Here's What Different Generations Say. Lexi Lane. People, April 18, 2026. https://people.com/different-generations-on-giving-up-dining-out-social-media-to-afford-concerts-11951944

members of Gen Z and Millennials have benefited from rising incomes, with Millennials experiencing a significant increase in household wealth. This increase has been attributed to better investment choices, as well as greater savings rates. In addition, Millennials and Gen Z have in some cases inherited significant amounts of money from older generations. Increased affluence has made it easier for younger consumers to finance significant expenditures on travel and live entertainment.

Reasons for More Extravagant Spending on Live Events

Among younger consumers there has been a noticeable trend toward "experiential" spending, where consumers are more interested in experiences than possession of physical goods. The satisfaction obtained from enjoyable memories, personal growth and increased connections can be greater than the consumption of physical goods. It can also provide an increased sense of well-being and fulfillment. Another factor that is increasing the demand for experiential events is known as the fear of missing out (FOMO.) Fear of missing out refers to the idea that an individual is missing out on important events or activities which can lead to dissatisfaction and anxiety. Fear of missing out can be a major determinant in the desire of young consumers to spend more on live events than on physical goods.

The rapid growth of social media such as Instagram, Tik Tok and others have increased the prevalence of FOMO as people see others enjoying activities and a lifestyle that they wish to share. Social media have also been able to promote live events in ways that were not possible with traditional forms of media.

Social Media and How They Promote Live Events

The rise of social media platforms, such as Instagram, Facebook, TikTok, YouTube, etc. has accelerated the trend toward experiential spending as it has become much easier to learn about experiences and events that a consumer can choose from, such as entertainment, sports, art, dining, etc. Social media has greatly increased the reach of marketing efforts compared with traditional methods such as advertising on TV. With Social Media such as Instagram and Facebook advertising can be specifically targeted to individual consumers based on the types of content that they look at on a regular basis.

Being able to view a huge variety of events can increase the fear of missing out (FOMO) experienced by consumers and thereby increase experiential spending. Social media also increases experiential spending by simplifying the process of paying for events.

Social media can also be used to improve marketing of concerts and festivals. For example, YouTube can be used to show videos of past concerts and other events to give fans an idea of what to expect at future events. A platform such as X can be used to post information about future concerts and festivals and provide feedback from past attendees. Hashtags can be used on platforms such as Instagram to enable fans to find posts about various concerts and other events. Pictures posted on these platforms can attract a huge number of followers. The organizers of the Glastonbury festival went out of their way to install a 4G mobile network to

make it easier for fans to post to their social media accounts during the festival. Social media can also make it easier for attendees to find out important details such as the schedule of events, the locations of different activities, etc.

Generational Differences in Consumer Behavior

Research has shown that there are several trends that can be observed in spending by younger consumers in 2026, including:

- Valuing experience over physical goods
- Using technology (such as the internet and mobile apps) to make purchases
- Being interested in health-related purchases
- Being willing to spend on educational materials
- Being willing to invest in cryptocurrencies as well as stocks
- Requiring more flexibility with payment arrangements

Research also shows a high comfort level with online shopping and the use of mobile apps for shopping among younger consumers. These indicate a trend toward more shopping on the internet through sites such as Amazon.com as well as more spending on experiential events rather than physical goods.

Research has shown that there are several reasons why online shopping is so satisfying to consumers. One of the most important reasons is convenience; it eliminates the need to travel to a store, to spend time looking for merchandise, etc. and avoids the risk of finding out that the store has run out of whatever product the consumer is looking for. These factors reduce the satisfaction that accompanies the purchase of a new good or service. Waiting time is eliminated when buying a product online. Searching for new products is greatly simplified especially due to recommendations by websites such as Amazon.com. Rapid delivery times also contribute to the satisfaction obtained from online shopping. The ease of returning unsatisfactory merchandise is another contributing factor to the popularity of online shopping. Overall, online shopping provides instant gratification, which is one of the main reasons that it is so popular.

The ultimate high-tech financial asset, cryptocurrencies, have proven to be highly popular with members of Gen Z and Millennials compared with older generations. This can be attributed to many factors, one of which is the novelty of a pure high tech financial asset compared with traditional investment assets such as stocks and bonds. Another attractive feature is the ability of investors to buy fractional units of a cryptocurrency, thereby reducing the cost and risk. Liquidity is another benefit to cryptocurrencies as they can be easily and quickly bought and sold. The fact that cryptocurrencies are not under the control of any government is another factor attracting young investors. Another potential factor is known as “FOMO” (fear of missing out) as young investors do not want to miss out on any potentially lucrative opportunities.

Gen Z and Millennials also require more payment flexibility than prior generations. They are far more likely to use technology such as P2P (“peer to peer”) and digital wallets to make

payments and are also much more likely to use a "buy now, pay later" approach to financing purchases than earlier generations. Gen Z and Millennials are also less likely to pay off credit card balances than earlier generations.

These trends have influenced spending in other areas such as travel and are also part of a larger trend toward more spending on experiences and less on physical goods. Research has shown that during the period from August 31, 2023 to August 31, 2024 spending by Americans on experiences grew by 32% compared with only 21% for physical goods.[110] In addition, spending on digital platforms such as Amazon.com (compared with in-person shopping) grew by about 188% from August 2019 to August 2024.

The value of the U.S. experiential market reached $55.53 billion in 2026 and is expected to reach $71.22 billion by 2035.[111] The global experiential market reached a new record $128.35 billion in 2024.

In addition, 25% of U.S. adults increased their experiential spending in 2025; they cited better value for their money than the purchase of physical goods. Their reasons include the creation of memories, relaxation, spending time with friends and family, etc. They also expressed an interest in trying new activities, spending time with a hobby, and getting more value for their money. One consequence of this is the trend toward giving more gifts based on experiences rather than physical goods.

Conclusion

In recent years members of two generations (Millennials and Gen Z) have shown a significant increase in their willingness to spend large sums of money on live events (such as concerts, festivals, sporting events, etc.) compared with previous generations (in particular, Baby Boomers and Generation X.). The younger generations have shown that their priorities are a bit different than previous generations. They are more interested in experiences than having physical goods and services and are willing to pay for the privilege. This occurs in spite of the fact that they are at a relatively early stage of their careers so that their incomes are typically lower than for older generations and are willing to make significant sacrifices to attend live events. In addition, the price of live events has increased significantly in inflation-adjusted terms in recent years. This phenomenon has been dubbed "funflation."

One important factor in this trend towards more spending on live entertainment is the increased ease of obtaining credit to these events compared with just a few years ago. This has come about due to improved financial technology and the high cost of live events. Vendors such as PayPal have found it necessary to offer generous financing terms to sell tickets to different events.

---

110 "The Shift to Digital and Experiential Spending Continues to Shape Consumer Trends", Get Your Guide. Sept. 30, 2024. https://www.getyourguide.press/blog/the-shift-to-digital-and-experiential-spending-continues-to-shape-consumer-trends

111 "Experiential Marketing Statistics 2026", Kande. https://www.kandephotobooths.com/blog/experiential-marketing-statistics/

There are several potential explanations for this change in consumer behavior among younger generations. One of these is the fact that these generations have always had access to the internet and for Gen Z, online social media and smart phones. These technological advances make it easier for people to find out about events that will take place in the future and increase the demand for these events and contribute to Fear of Missing Out (FOMO.) The generations that are constantly exposed to social media are aware that if they do not attend live events, they will easily find out what they have missed and may not want to risk feeling that they missed out on a golden opportunity that may never be available again. They may also not want others to find out that they have missed these events, which is much likelier since the rise of social media. This trend toward spending progressively more on live events is not expected to slow anytime soon. Several financial institutions are specifically aiming to increase the number of customers among Millennials and Gen Z by offering them credit cards and other financial services that reward them for spending on live experiences (such as restaurants, travel, concerts, etc.) instead of physical goods and services. The amount of consumer spending on these live activities is increasing a percentage of total spending and is likely to continue rising in the near future. It is possible that the youngest generations (Generation Alpha and Generation Beta) will find themselves spending even more of their incomes on these types of activities, possibly driven by changes at financial institutions and continuing advances in technology. Several financial institutions are attempting to increase their market share by increasingly catering to the market for younger consumers and experiential spending.

Appendix: Generation Names

The names of the generations dating back to the late nineteenth century are generally agreed to be:[112]

The Lost Generation – 1883 - 1900
The Greatest Generation – 1901 – 1927
The Silent Generation – 1928 – 1945
Baby Boomers – 1946 – 1964
Generation X – 1965 – 1980
Millennials (also known as Generation Y) – 1981 – 1996
Generation Z – 1997 – 2010
Generation Alpha – 2011 – 2024
Generation Beta – 2025 – 2039

Many of these names were first introduced by authors Neil Howe and William Strauss in their 1991 book "Generations." The Lost Generation name is often attributed to the author Gertrude Stein.

[112] "A Year-by-Year Guide to the Different Generations". Parents. Sarah Cottrell, September 26, 2025.

Generation Alpha and Beta would be too young to buy concert tickets or attend other events that are intended for adults. Therefore, the two youngest generations of adults would be considered the Millennials and Generation Z. The Millennials are the first generation to see the introduction of the internet; they are also the largest of the generations, totaling 74.8 million people or 21.8% of the population. Generation Z is the first generation to have access to the internet, smart phones and tablets during their entire lifetimes.

Works Cited

Coachella. https://www.google.com/search?q=coachella&oq=coachella&gs_lcrp=EgZjaHJvbWUyDAgAEEUYORixAxiABDIKCAEQABixAxiABDIKCAIQABixAxiABDIKCAMQABixAxiABDINCAQQABiDARixAxiABDIHCAUQABiABDIHCAYQABiABDIKCAcQABixAxiABDIKCAgQABixAxiABDIHCAkQABiABNIBCDIwODVqMGo3qAIAsAIA&sourceid=chrome&ie=UTF-8

Google What Were Ticket Prices for Coachella 2026 https://www.google.com/search?q=what+were+tickets+prices+for+coachella+2026&oq=what+were+tickets+prices+for+coachella+2026&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIICAEQABgWGB4yCAgCEAAYFhgeMg0IAxAAGIYDGIAEGIoFMg0IBBAAGIYDGIAEGIoFMg0IBRAAGIYDGIAEGIoFMg0IBhAAGIYDGIAEGIoFMgoIBxAAGIAEGKIE0gEINzM5MmowajeoAgCwAgA&sourceid=chrome&ie=UTF-8

LA Times What People Spent on Coachella https://www.latimes.com/entertainment-arts/music/story/2026-04-22/people-tell-us-how-much-they-spent-on-coachella-2026

Are There Other Expensive Music Festivals Such as Coachella? https://www.google.com/search?q=are+there+other+expensive+music+festivals+such+as+coachella&oq=are+there+other+expensive+music+festivals+such+as+coachella&gs_lcrp=EgZjaHJvbWUyCQgAEEUYORigATIHCAEQIRigATIHCAIQIRigAdIBCTEzMzUwajBqN6gCALACAA&sourceid=chrome&ie=UTF-8

Digital Music News. "The 10 Most Expensive Music Festivals in the World." https://www.digitalmusicnews.com/2017/07/04/expensive-best-music-festivals-world/

What is the Average Age and Income of the People Who Attend Coachella? https://www.google.com/search?q=what+is+the+average+age+and+income+of+people+who+attend+coachella&oq=what+is+the+average+age+and+income+of+people+who+attend+coachella&gs_lcrp=EgZjaHJvbWUyBggAEEUYOdIBCDg2ODlqMGo3qAIAsAIA&sourceid=chrome&ie=UTF-8

"How to Survive Becoming a Coachella Elder." Vogue. April 18, 2024. Alana Hope Levinson. https://www.vogue.com/article/coachella-elder#:~:text=When%20I%20went%20to%20Coachella,can%20still%20get%20press%20passes.)

"16 Coachella Attendees Revealed Their Job, Yearly Salary and Whether They Regretted Their $600 Ticket". https://www.buzzfeed.com/pernellquilon/coachella-2023-real-experiences

Research on Why People are Willing to Spend so Much on Concerts and Festivals https://www.google.com/search?q=research+on+why+people+are+willing+to+spend+so+much+on+concerts+and+festivals&oq=research+on+why+people+are+willing+to+spend+so+much+on+concerts+and+festivals&gs_lcrp=EgZjaHJvbWUyBggAEEUYOdIBCTEyOTUxajBqN6gCALACAA&sourceid=chrome&ie=UTF-8

Tech, Media and Telecom. "Funflation" – and Live Event Hype – Goes up Against More Cost-Conscious Consumers". October 8, 2024. Jana Arbanas, et al.

https://www.deloitte.com/us/en/insights/industry/technology/digital-media-trends-consumption-habits-survey/2024/funflation-goes-up-against-cost-conscious-consumers.html
Lending Tree. "Fans Shake Off High Ticket Prices for the Love of Rock and Roll". https://www.lendingtree.com/credit-cards/study/ticket-prices-concerts/
"Would You Give up Dining Out or Delete Social Media to Afford Concert Tickets? Here's What Different Generations Say. Lexi Lane. People, April 18, 2026. https://people.com/different-generations-on-giving-up-dining-out-social-media-to-afford-concerts-11951944
"Here's How Much Millennials and Gen Z Are Spending on Concerts". Ophelia Garcia Lawler. Thrillist. April 18, 2024. https://www.thrillist.com/news/nation/millennial-spending-concerts-live-events-travel
"Going the Extra Mile: Three in Five Younger Travelers Venture out of State For Live Events." Bread Financial, April 16, 2024. https://newsroom.breadfinancial.com/aaa-summer-travel-survey
Gen Z Spending Behaviors. Merge. September 1, 2024. https://www.mergeworld.com/insights/consumer/gen-z-spending-behaviors
How Experiential Spending Can Positively Affect Your Life. Marketing. May 13, 2024. https://www.kentwood.com/blog/2024/05/13/how-experiential-spending-can-positively-impact-your-life/
Millennial Spending Habits: 7 Trends Marketers Need to Know in 2026. GWI. Caitlin Nuttall. https://www.gwi.com/blog/millennial-spending-habits
Why Online Shopping Makes You So Happy. Time. Angela Haupt. July 26, 2022. https://time.com/6200717/online-shopping-psychology-explained/
"How Coachella Set the Standard for Modern Music Festivals". Push Agency. May 12, 2025. https://pushmodels.com/blog/coachella/
"The Surprising Reason Why Millennial Wealth is Growing Faster Than Other Generations". Investopedia. Karen Kroll, May 23, 2025. https://www.investopedia.com/millennial-wealth-growth-11718919
"The Shift to Digital and Experiential Spending Continues to Shape Consumer Trends", Get Your Guide. Sept. 30, 2024. https://www.getyourguide.press/blog/the-shift-to-digital-and-experiential-spending-continues-to-shape-consumer-trends
"Experiential Marketing Statistics 2026", Kande. https://www.kandephotobooths.com/blog/experiential-marketing-statistics/
"The Experience Economy is Resurgent, But Only Among Select Consumers". Intelligence. Lindsey Roeschke, August 26, 2025. https://intel.morningconsult.com/mc-content/analysis/experience-economy-consumer-research
"How Gen Z and Rich Consumers are Re-Shaping the Experience Economy." Forbes. Jeff Fromm, November 9, 2023.

https://www.forbes.com/sites/jefffromm/2023/11/09/gen-z-and-affluent-consumers-are-re-shaping-the-experience-economy/

"A Year-by-Year Guide to the Different Generations". Parents. Sarah Cottrell, September 26, 2025. https://www.parents.com/parenting/better-parenting/style/generation-names-and-years-a-cheat-sheet-for-parents/

"Some Concertgoers Use Buy Now, Pay Later Loans for Tickets. Start a 'Beyonce Fund' Instead, Analyst Says. CNBC.com, August 9, 2025. Ana Teresa Sola. https://www.cnbc.com/2025/08/09/using-buy-now-pay-later-bnpl-loans-for-concert-tickets.html

"TED: The Economics Daily", Bureau of Labor Statistics. February 5, 2026. https://www.bls.gov/opub/ted/2026/big-games-big-prices-admissions-for-sporting-events-up-123-percent-since-2000.htm

"American Express, Chase and Capital One Seek Gen Z Live Event Junkies". Forbes, September 12, 2025. by Jeff From. https://www.forbes.com/sites/jefffromm/2025/09/12/american-express-chase-and-capital-one-seek-gen-z-live-event-junkies/

U.S. Bureau of Labor Statistics. "Big Games, Big prices: Admissions for Sporting Events up 123 Percent Since 2000". February 5, 2026. https://www.bls.gov/opub/ted/2026/big-games-big-prices-admissions-for-sporting-events-up-123-percent-since-2000.htm

Is an Unenforced Law Really a Law? By David Liang

Abstract

In exploring whether an unenforced law maintains its validity, this paper starts with a hypothetical scenario involving a prescribed bedtime and parental punishment. After defining punishment in a legal context and establishing assumptions regarding the given scenario, the paper argues that a law without enforcement is impotent. This claim is substantiated through three lenses: legal philosophy, judicial practice and empirical evidence. Drawing from Alexander Hamilton's *Federalist No.15,* as well as natural law and normative ethics theories, the paper illustrates how enforcement is a defining characteristic of legal order. In practical jurisprudence, the doctrine of *desuetude* recognizes that prolonged non-enforcement renders a statute null and void. Lastly, data from a field experiment demonstrate that unenforced regulations fail to alter behavior and instead inflict societal and psychological costs. The paper concludes that legal legitimacy is inextricably tied to enforcement; without consistent penalties for disobedience, institutional authority erodes, and the law itself ceases to exist as a binding legal order.

Introduction

Suppose your parents say that 11pm is your bedtime, but they don't punish you if you don't go to bed by 11pm. Is 11pm really your bedtime?

The hypothetical described in the prompt brings up important questions about the validity and effectiveness of laws that lack enforcement. This broad and complex issue can take the form of various scenarios, but at its core, it challenges the law's relevance in the absence of punishment.

To answer the question, it's important to first establish key assumptions. In criminal law, punishment is defined as any pain, penalty, suffering, or confinement inflicted upon a person by the authority of the law and the judgment and sentence of a court, for some crime or offense committed by him, or for his omission of a duty enjoined by law. (PUNISHMENT Definition & Meaning, n.d.)

From the definition, we can identify three essential components of punishment. One, punishment is an infliction of some pain or loss on the person or group who violates the law; two, it is imposed by a recognized authority; and three, it is intended to remedy and deter the wrongdoing. Due to its deterrent nature, punishment serves as an effective tool in law enforcement. *Lack of punishment in the given scenario can be characterized as, or equated to, a lack of enforcement.*

The essay prompt is short on details of the scenario. For the purpose of discussion, I will assume that *no punishment was given at all for long periods* for violation of the bedtime rule. Sporadic and inconsistent punishment is beyond the scope of this essay, as it pertains to a very different legal setting.

This paper will argue that the bedtime is not 11pm, or, in a legal context, a law without enforcement is impotent. My argument is based on the following three principles: (1) according to legal philosophy, a law must include an enforcement mechanism or it is not a true law; (2) in

the world of legal practice, “desuetude” deems a law null and void after long periods of non-use; and (3) empirically, a law without enforcement loses its social impact and even generates costs.

Legal Philosophy

a. Federalist 15

Federalist 15 argues that law without the power of enforcement is inherently flawed. The government merits authority to create laws that establish standards for its citizens. Through its separate branches, the government must work together to validate and preserve these laws. It is the duty of the government to use its power, inflicted through either the court or military, to ensure a law-abiding population. Law without punishment can be perceived as unreputable, or unable to repress acts of disobedience. Therefore, law without the necessary sanctions can be seen as merely advice. (Hamilton, Alexander, 1787)

Federalist 15 emphasizes this principle through its analysis of the defects of the Articles of Confederation. The core issue with the Articles was lack of a coherent federal government. Hamilton criticizes the Articles of Confederation for being inadequate in establishing a secure justice system, paying debt, and gathering troops. The Articles of Confederation placed much of its power on the states, and was unable to regulate them, causing the states to completely disregard the law. States viewed themselves as sovereign, rather than as part of one nation united under the Articles of Confederation. Hamilton feared a future in which the country disregarded its government and the will of the people, warning others of an “impending anarchy” that required immediate prevention. Federalist 15 deftly analyzes the flaws and consequences of a government that does not enforce its laws.

The parents' authority over the children in the prompt is akin to the Articles of Confederation. The parents fail to apply consistent punishments to enforce their rules. Similar to the states in the decentralized system of the Articles of Confederation, which struggled to unite under a single federal government, the children of the family may see themselves as sovereign entities due to the lack of enforcement, causing them to believe that they are beyond the consequences of parental authority. This effect caused by the lack of discipline can disrupt the structure of the family, with the children neglecting the power of the parents. The 11pm bedroom policy is therefore not a rule to be followed, but a mere recommendation.

b. Natural Law and Normative Ethics

The necessity of enforcement as a principle of law is also widely acknowledged in natural law theories and normative ethics. Both recognize that deterrence plays a pivotal role in the justification of law. Just punishment must be given to those who break the law to not only punish, but to guide one away from disobedience. The principle of punishment is to restore balance to order that has been disrupted, and to ensure the validity of the law among the people. Deterrence is the goal of punishment, aiming to drive the offender and others away from committing the same offenses. Punishment must be active in a society that wishes to utilize

deterrence. In order for deterrence to be successful, the consequences of breaking the law must be credible, fair and severe enough to encourage obedience. (Finnis, John, 2007)

Certain social objects are built upon core values, and when these values are not found, the objects diminish, or cease to exist. These objects are described as having a "normative constitution." An example of this is the idea of a family. The core principle of a family is that each member loves and cares about one another. Other attributes that make up a family, such as a group living together or being biologically related, are unified under mutual love. When mutual love is absent, the "family" can be considered not a family at all. Law also has a normative constitution. While many other properties and values contribute to the core of law, such as its enactment and codification, they are fundamentally bound by the power of enforcement. Participation in a society requires obedience to the law, and this acceptance of the law stems from enforcement, in that the law must generate enough respect among citizens to deserve to be followed. The objective is not that one must simply follow the law, but that one ought to follow the law out of either respect or fear of the government. The value of law is constituted on enforcement. Without it, a law can't instantiate itself being a legal order. Consequently, the 11pm rule is only normative, but not a law, without punishment. (Kleinfeld, Joshua, 2011)

Legal Practice

Desuetude is a concept that renders a law invalid after long periods of disuse. United States vs Elliot is a prime example of the issue that arises when law fails to be enforced. Jay Elliot and other conspirators, the defendants, planned to destroy a railroad bridge in the Republic of Zambia in order to halt the supply of Zambian copper in the global market. The defendants planned to then sell their supply of copper following the copper shortage. These acts violated the 18 U.S.C § 956, which states that if two or more citizens of the United States conspire to destroy or deface federal property within a foreign country, that each of the members would be fined no more than $5000 or imprisoned for 3 years or less. The question of the case is whether or not the convictions of Elliot and his fellow conspirators under the RICO statute could be upheld. (United States v. Elliott | Case Brief for Law Students, n.d.)

However, the core issue became one that extended beyond the original scope of the case, in that the U.S. did not enforce the statute in the past, which caused confusion for the defendant. While the court ruled against the defendant in his argument employing desuetude, the judge's opinion explained the scenarios in which desuetude could be applied. One, desuetude may only be relevant when a statute is still applicable. Otherwise, the statute may be obsolete already and non-enforcement therefore is not a factor. Two, a statute is in desuetude status when violation has become customary through a long period of non-enforcement. And lastly, a statute in a state of desuetude creates serious issues of fair notice and potential for abuse with selective enforcement that may violate equal protection doctrines. (United States v. Elliott, 1967)

Following the principles laid out in the United States vs. Elliot, we can see that desuetude is applicable to the scenario depicted in the essay prompt. There's no reason to believe that regulation on bedtime is outdated. And it is safe to assume that violation of the rule has become

the norm after a long period of non-punishment. Additionally, a lack of enforcement and any other indication that 11pm is the established bedtime infers absent notice and a potential ill-use of selective enforcement. Therefore, the 11pm bedtime rule is no longer practical.

Empirical Evidence

It's often observed in the real world that a law loses its power when there is no enforcement. The effects of non-enforced law are not only limited to undesired behavior of breaking the law, but also include more profound detriments to society. Common outcomes include undermined respect for the law, eroded authority, and hidden psychological costs.

An NYU Law School article explored the psychological costs of unenforced law in a public environment. Professors Ben Depoorter and Stephan Tontrup conducted an experiment in a German bar, where they observed the impact when non-smoking rules were not enforced. The bar designated one room with alternating smoking policies: sometimes the bar employed the smoking allowed rule, while other times the bar assigned the room as smoking prohibited but did not enforce the prohibition. The study compared the behavior of the customers under these different policies. The results showed that customers would spend less money and a shorter amount of time in the bar when the room was designated as smoking-prohibited but not enforced. The experiment demonstrated that customers would exhibit these behaviors because they felt an infringement of their rights to a smoking-free environment even if they were indifferent to the smoke exposure. Although the customers did not suffer from any further physical harm, the violation of no-smoking policy caused the shortened stay or smaller bill. (Depoorter, Ben, and Tontrup, Stephan, 2016)

Based on this experiment, we can infer that when there is no punishment for staying up past 11pm, other family members may suffer psychological costs. When siblings observe the rule being broken, they may feel their entitlement to peaceful sleep is violated, even if the siblings are indifferent about their bedtime. While the purpose of the rule is to encourage family members to enhance the well-being of themselves and others, it ultimately fails to fulfill its purpose, and even deteriorates the welfare of the family as a whole. Eventually, the rule will lose its legitimacy.

Conclusion

When punishment is absent, the 11pm rule loses its efficacy, undermining the fundamental value of the rule and raising questions about whether it is merely a recommendation. After long periods of non-enforcement, the rule becomes practically null, as it is difficult and even unfair to enforce it again. Staying up after 11pm may become the norm, rendering the rule completely irrelevant.

Society fails without enforced laws. Parents, as the authority within the family, are responsible for upholding its structure and establishing rules to ensure everyone's well-being, which justifies the bedtime regulation. However, when punishment is not given for breaking the rule, the consequences affect not only the children but the entire family. As the parents' authority

erodes, other family rules may be jeopardized, leading the family to stop functioning effectively. Consequently, not only does the 11pm bedtime rule become void, but its validity is also questioned as the family begins to disintegrate.

Works Cited

Black's Law Dictionary. “PUNISHMENT Definition & Meaning.” The Law Dictionary. https://thelawdictionary.org/punishment/. Accessed June 25, 2024.

Depoorter, Ben, and Stephan Tontrup. "The Costs of Unenforced Laws: A Field Experiment." 2016.

Finnis, John. “Natural Law Theories.” Stanford Encyclopedia of Philosophy, 2007, https://plato.stanford.edu/entries/natural-law-theories/#ResPun.

Hamilton, Alexander. “The Avalon Project : The Federalist Papers : No. 15.” The Avalon Project, 1787, https://avalon.law.yale.edu/18th_century/fed15.asp.

Kleinfeld, Joshua. “Forum: Enforcement and the Concept of Law.” The Yale Law Journal, 2011, https://www.yalelawjournal.org/forum/enforcement-and-the-concept-of-law.

United States v. Elliott, 266 F. Supp. 318 (S.D.N.Y. 1967) “United States v. Elliott | Case Brief for Law Students.” Casebriefs, https://www.casebriefs.com/blog/law/criminal-law/criminal-law-keyed-to-dressler/inchoate-offenses/united-states-v-elliott/. Accessed June 25, 2024.

Investigation on Temperature-dependent Mechanical Properties of Several Typical Polymers

By Kexin Deng[1], Jiaheng Li[2], Derek W. Davies[3] and Junhui Hu[2]

[1] *Santa Margarita Catholic High School, 22062 Antonio Pkwy, Rancho Santa Margarita, CA 92688, America*
[2] *Shenzhen Baroy New Material Technology Co., Ltd., Shenzhen, 518000, China*
[3] *Department of Material Science and Engineering, University of Texas Austin, 110 Inner Campus Drive, Austin, TX 78705, America*
**Correspondence: E-mail: cathydeng829@gmail.com; derekdavies@utexas.edu*

Abstract: The mechanical properties of polymeric materials are known to be significantly influenced by temperature. In this work, the mechanical properties (ultimate tensile strength, UTS; ultimate flexural strength, UFS; Izod notched impact strength, INIS) of five typical polymers (LLDPE, PP-H, PMMA, TPU and EP) are investigated across a range of temperatures (-60, -20, 0, 25, 50 and 80℃) below the glass transition temperatures of these materials. The results showed that the mechanical properties of all materials are, to some extent, temperature-dependent. In terms of tensile properties, the UTS of LLDPE, PP-H, PMMA and EP decreased with increasing temperature. TPU was not fractured within the stroke of the testing machine, and its maximum tensile stress during stretching did not change much with temperature. With regard to the bending test results, only EP was fractured after testing. Its UFS decreased first and then increased when the test temperature increased, reaching the minimum value of 102.4 MPa at 0℃. As for the impact tests, ductile PP-H and TPU were not fractured, and the rest three materials were fractured after testing. As brittle materials, EP and PMMA exhibited very low INISs at all testing temperatures. The INIS of LLDPE demonstrated a negative correlation to temperature.

Keywords: Linear low density polyethylene (LLDPE) Homo-polypropylene (PP-H) Polymethylmethacrylate (PMMA) High-elastic polyurethane (TPU) Epoxy plastic (EP) Injection molding Hot-setting Temperature Mechanical properties.

1. Introduction

Polymeric materials have undergone a great development over the past century and have become indispensable components in today's society [1]. They are extensively used in the application domains such as biomedical [2], construction [3], energy [4] and electronics [5-7] due to their advantages of low cost, low density, good corrosion resistance, good insulation and excellent processability, etc. So far, the production volumes of commodity polymeric materials (e.g., polyethylene and polystyrene) have surpassed those of traditional metallic materials [8]. Mechanical properties (e.g., tensile properties, compressive properties and bending properties) are basic considerations for the engineering application of polymers. Therefore, it is meaningful to carry out research on the mechanical properties of polymers.

Polymers are usually exposed to different surrounding temperatures and their mechanical properties are known to be significantly influenced by temperature. Li et al. [9] investigated the temperature dependence of the mechanical modulus of a polymer blend with 70wt % linear low density polyethylene (LLDPE) and 30wt % high density polyethylene (HDPE). It was found that the mechanical modulus of the blend decreased with increasing temperature because the thermal motion of polyethylene molecules increases with rising temperature. Ref. [10] conducted flexural and compressive tests at different temperature levels for epoxy and unsaturated polymer mortars. The results showed that a significant loss in flexural/compressive strength is observed upon increasing the test temperature for both polymer mortars. Richeton et al. [11] carried out compressive tests for directional polymethylmethacrylate (PMMA) at temperatures ranging from -40℃ to the glass transition temperature (Tg). Their results showed that Young's modulus, yield strength, and the strain-hardening rate decreased when the temperature increased. Gao et al. [12] revealed the tension-compression asymmetry of directional PMMA. Hu et al. [13] conducted compressive tests on directional PMMA from -70 to 120℃ and proposed a thermo-viscoelastic constitutive model. Adel et al. [14] conducted three-point bending tests and uniaxial tensile tests for directional PMMA from 20 to 80℃. Ashrafizadeh and co-workers [15] tested the tensile properties of several high-elastic polyurethanes (TPUs) at 22, 60 and 100℃ for strains up to 350%. They found that all the PU elastomers became softer with increasing temperature. Tai et al. [16] evaluated the effect of temperature on the impact fracture behaviour for the PP/PE blends via the Charpy and drop weight plate impact tests. During their tests, impact strengths reduced with decreasing testing temperature. To date, few studies focus on the temperature dependence of multiple mechanical properties of various polymeric materials.

In this work, the mechanical properties (ultimate tensile strength, UTS; ultimate flexural strength, UFS; Izod notched impact strength, INIS) of four typical thermoplastics (LLDPE; homo-polypropylene, PP-H; PMMA and TPU) and one thermoset (epoxy plastic, EP) are investigated across a range of temperatures (-60, -20, 0, 25, 50 and 80℃).

The research results are useful to the engineering applications of the five polymeric materials.

2. Materials and Methods

2.1 Sample Preparation

LLDPE/PP-H/PMMA/TPU samples were prepared by Shenzhen Hualixing New Materials Co., Ltd. using injection molding (IM) technology. EP samples were prepared by Shenzhen Baroy New Material Technology Co., Ltd. using a thermosetting method. The nominal geometric dimensions of the mechanical property testing samples are illustrated in Fig. 1. For injection-molded samples, the IM direction is consistent with their length direction.

2.2 Mechanical Properties Tests

The sample design and implementation of uniaxial quasi-static tensile, three-point bending and Izod notched impact testing were according to the ISO 527-2:2025, ISO 178: 2010 and ISO 180: 2023 standards, respectively.

A SANS/CMT 6104 universal material testing machine was used for tensile and bending property tests. A ZBC1400-B Izod impact test machine was used for Izod notched impact testing. A constant temperature and humidity chamber (BELL/BE-TH-150L8) was used to heat or cool the samples. The planned tests were carried out at six different application-relevant temperatures (-60, -20, 0, 25, 50 and 80℃). For the tensile and bending tests, the whole testing process was conducted in the environmental chamber, which was heated and cooled using resistance wire and liquid nitrogen, respectively. For the impact tests, the samples were placed in the environmental chamber for 15 min to reach the target temperature and then taken out quickly for testing. Since the duration of one impact test did not exceed 15 seconds, it could be considered that the temperature did not change significantly during the test. The tensile stretching rate was fixed at 5 mm/min. For the bending test, the loading rate was 2mm/min and the span length was 64 mm. Eight samples per condition were tested, and the mean values were taken for the analysis.

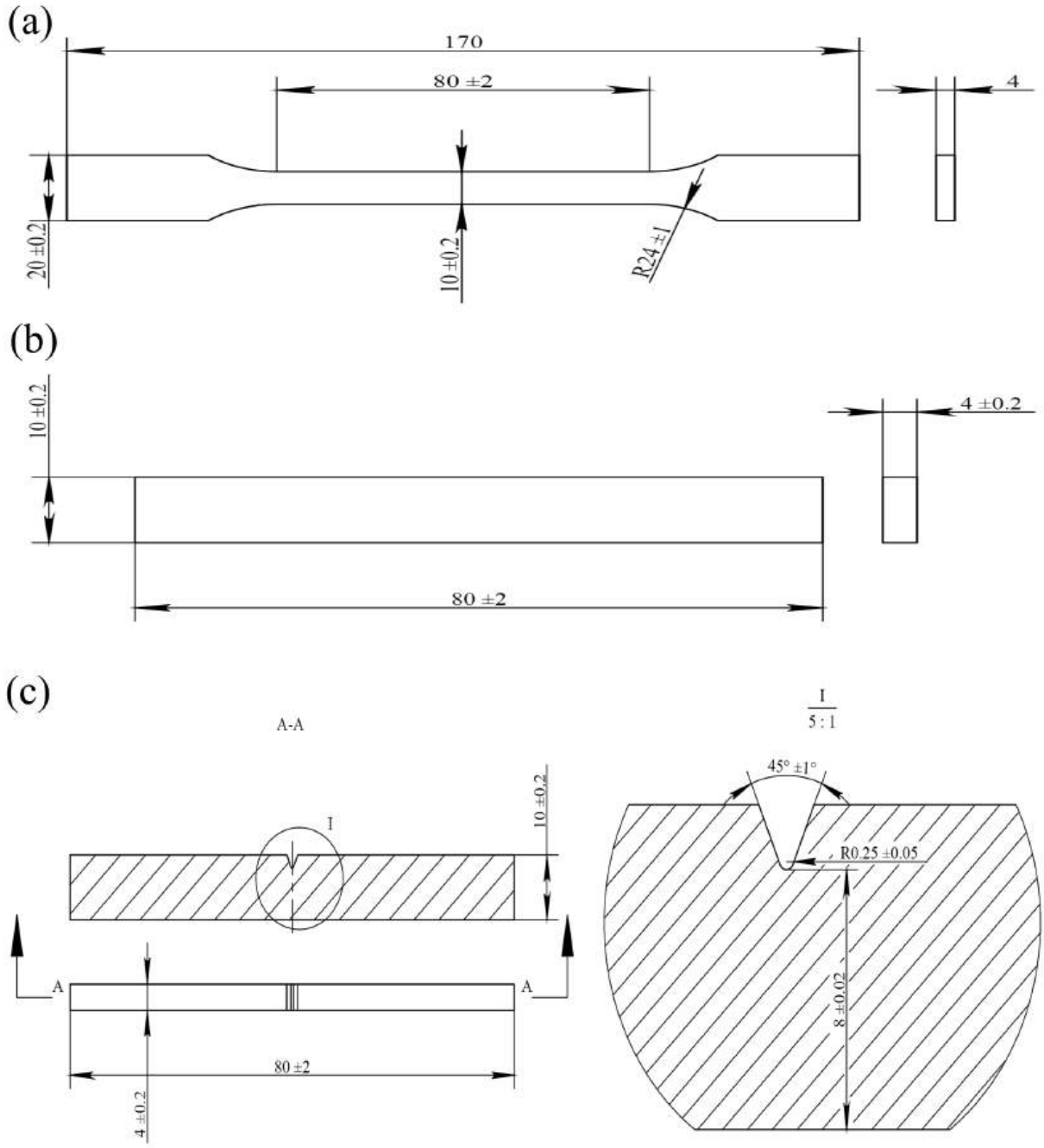


Fig. 1 Size diagrams of the mechanical property testing samples: (**a**) uniaxial tensile, (**b**) three-point bending, and (**c**) Izod notched impact.

2.3 Determination of Glass Transition Temperature

The Tgs of the investigated polymeric materials were determined by the TMA method according to the ISO 11359-2:2021 standard. Cubic samples (5 mm x 5 mm x 5 mm) were heated from 40℃ with a constant heating rate of 10℃/mim in the atmosphere of nitrogen (the gas flow rate was set as 40 mL/min) under a compression force of 0.05 N.

3. Results and Discussion

3.1 Temperature Dependence of Mechanical Properties

Fig. 2 gives the variation of the UTS, UFS and INIS of the five polymers with the change of temperature. It can be seen that the mechanical properties of all materials are more or less temperature-dependent. The UTS of LLDPE, PP-H, PMMA and EP decreased with increasing temperature (i.e., these materials exhibited thermal softening behavior), as illustrated in Fig. 2a. Generally, the decrease in the UTS of polymers with increasing temperature can be attributed to the enhancement of the movement ability of molecular chains. TPU was not fractured within the stroke of the testing machine, and its maximum tensile stress during stretching did not change much with temperature. The fractured samples failed randomly in the gauge length. LLDPE, PP-H showed ductile fracture characteristics with considerable necking formed. While PMMA and EP exhibited brittle fracture characteristics, and a wave-like fracture morphology with no significant necking was observed. Also, the stress whitening phenomena were observed in the gauge distance of the tensile samples for LLDPE and PP-H after stretching.

Bending tests showed that only EP was fractured after testing and the rest four materials were not fractured within the stroke of the testing machine (see Fig. 2b). The UFS of the EP decreased first and then increased when the test temperature increased, reaching the minimum value of 102.4 MPa at 0℃. As for the four unfractured materials, the UFS of TPU and LLDPE almost remained unchanged as temperature varied; the UFS of PP-H decreased slightly with increasing temperature; and the UFS of PMMA increased first and then decreased when the test temperature increased, reaching the maximum value of 114.0 MPa at 25℃.

As shown in Fig. 2c, PP-H and TPU were not fractured, and the rest three materials were fractured after testing. As brittle materials, EP and PMMA exhibited very low INISs at all testing temperatures. In addition, the INIS of PP-H exhibited a positive correlation to temperature, while that of LLDPE and TPU demonstrated a negative correlation to temperature.

The average values of the UTS, UFS and INIS at each temperature for each material were calculated, as listed in Table 1, Table 2 and Table 3, respectively. At all temperatures, the order of UTS from high to low is: PMMA, EP, PP-H and LLDPE. And the order of UFS from high to low is: EP, PMMA, PP-H, LLDPE and TPU.

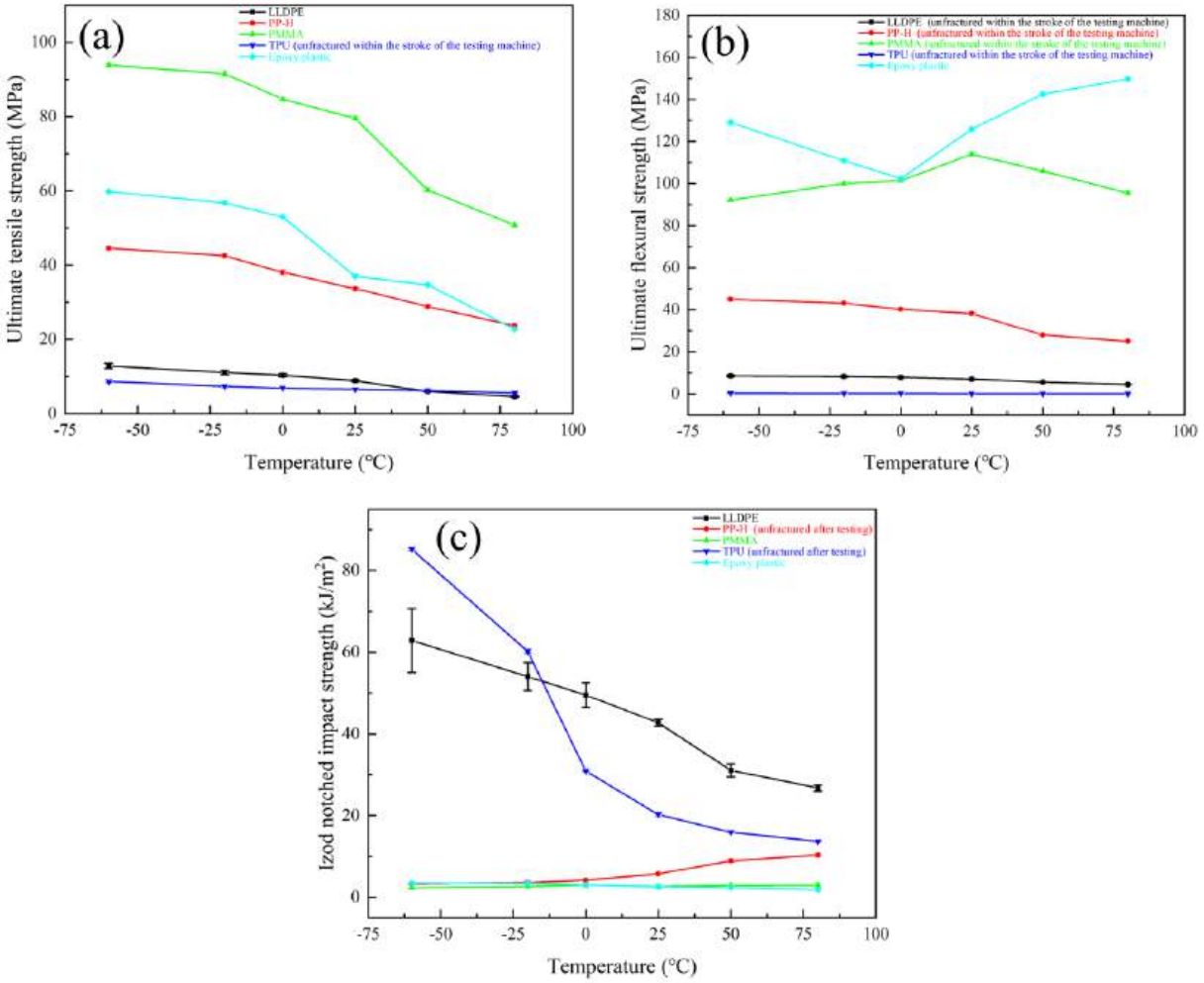


Fig. 2 Temperature dependence of mechanical properties: (**a**) uniaxial tensile, (**b**) three-point bending, and (**c**) Izod notched impact.

Table 1 Ultimate tensile strength of the five investigated polymers at different test temperatures.

Test temperature (℃)	Ultimate tensile strength of different materials (MPa)				
	LLDPE	**PP-H**	**PMMA**	**TPU**	**EP**
-60	12.9 ± 0.7	44.6 ± 0.5	93.8 ± 0.4	8.7 ± 0.3	59.8 ± 0.2
-20	11.0 ± 0.6	42.5 ± 0.4	91.5 ± 0.8	7.3 ± 0.3	56.8 ± 0.3
0	10.3 ± 0.4	38.0 ± 0.7	84.7 ± 0.3	6.9 ± 0.1	53.0 ± 0.1
25	8.9 ± 0.2	33.6 ± 0.5	79.5 ± 0.5	6.5 ± 0.2	36.9 ± 0.4
50	6.0 ± 0.1	28.9 ± 0.3	60.2 ± 0.4	6.2 ± 0.1	34.7 ± 0.3
80	4.6 ± 0.2	23.7 ± 0.5	50.8 ± 0.4	5.5 ± 0.2	22.7 ± 0.5

Notes: The TPU samples were not fractured within the stroke of the testing machine and the maximum engineering stress values during testing were reported as ultimate tensile strengths.

Table 2 Ultimate flexural strength of the five investigated polymers at different test temperatures.

Test temperature (℃)	Ultimate flexural strength of different materials (MPa)				
	LLDPE	PP-H	PMMA	TPU	EP
-60	8.6 ± 0.3	45.0 ± 0.5	92.2 ± 0.	0.3 ± 0.1	129.1 ± 1.3
-20	8.3 ± 0.2	43.1 ± 0.3	100.1 ± 0.	0.2 ± 0.0	110.9 ± 0.9
0	7.8 ± 0.1	40.2 ± 0.5	101.5 ± 0.	0.2 ± 0.0	102.4 ± 0.7
25	7.0 ± 0.1	38.2 ± 0.4	114.0 ± 0.	0.1 ± 0.0	125.9 ± 0.9
50	5.7 ± 0.1	28.0 ± 0.3	105.9 ± 0.	0.1 ± 0.0	142.5 ± 1.0
80	4.5 ± 0.1	25.0 ± 0.4	95.5 ± 0.8	0.1 ± 0.0	149.7 ± 0.8

Notes: The LLDPE, PP-H, PMMA and TPU samples were not fractured within the stroke of the testing machine and the maximum engineering stress values during testing were reported as ultimate flexural strengths.

Table 3 Izod notched impact strength of the five investigated polymers at different test temperatures.

Test temperature (℃	Izod notched impact strength of different materials (KJ/m²				
	LLDPE	PP-H	PMMA	TPU	EP
-60	62.9 ± 7.8	3.4 ± 0.1	2.3 ± 0.1	85.2 ± 0.4	3.5 ± 0.0
-20	54.0 ± 3.4	3.6 ± 0.3	2.6 ± 0.4	60.2 ± 0.5	3.4 ± 0.0
0	49.5 ± 3.0	4.2 ± 0.1	3.0 ± 0.6	30.9 ± 0.2	3.0 ± 0.0
25	42.8 ± 0.8	5.7 ± 0.2	2.6 ± 0.6	20.3 ± 0.5	2.6 ± 0.1
50	31.1 ± 1.6	8.9 ± 0.4	2.9 ± 0.4	16.0 ± 0.2	2.4 ± 0.0
80	26.8 ± 0.7	10.3 ± 0.3	3.0 ± 0.7	13.7 ± 0.3	1.8 ± 0.1

Notes: The PP-H and TPU samples were not fractured after testing.

3.2 TMA curves

The glass transition is an inherent property of amorphous polymeric materials and a macroscopic manifestation of the transformation of polymer motion forms. It directly affects the service performance and processing performance of polymeric materials. Taking Tg as the boundary, polymeric materials have different physical properties. In the plastic industry, Tg is generally regarded as the upper limit of the service temperature of plastic products and the lower limit of that of elastomers. The Tgs of LLDPE, PP-H, PMMA, TPU and EP were tested by TMA as 122.26, 162.81, 107.84, 122.12 and 157.83℃, *respectively* (see Fig. 3). That is, for all materials, the testing temperatures are below their Tgs.

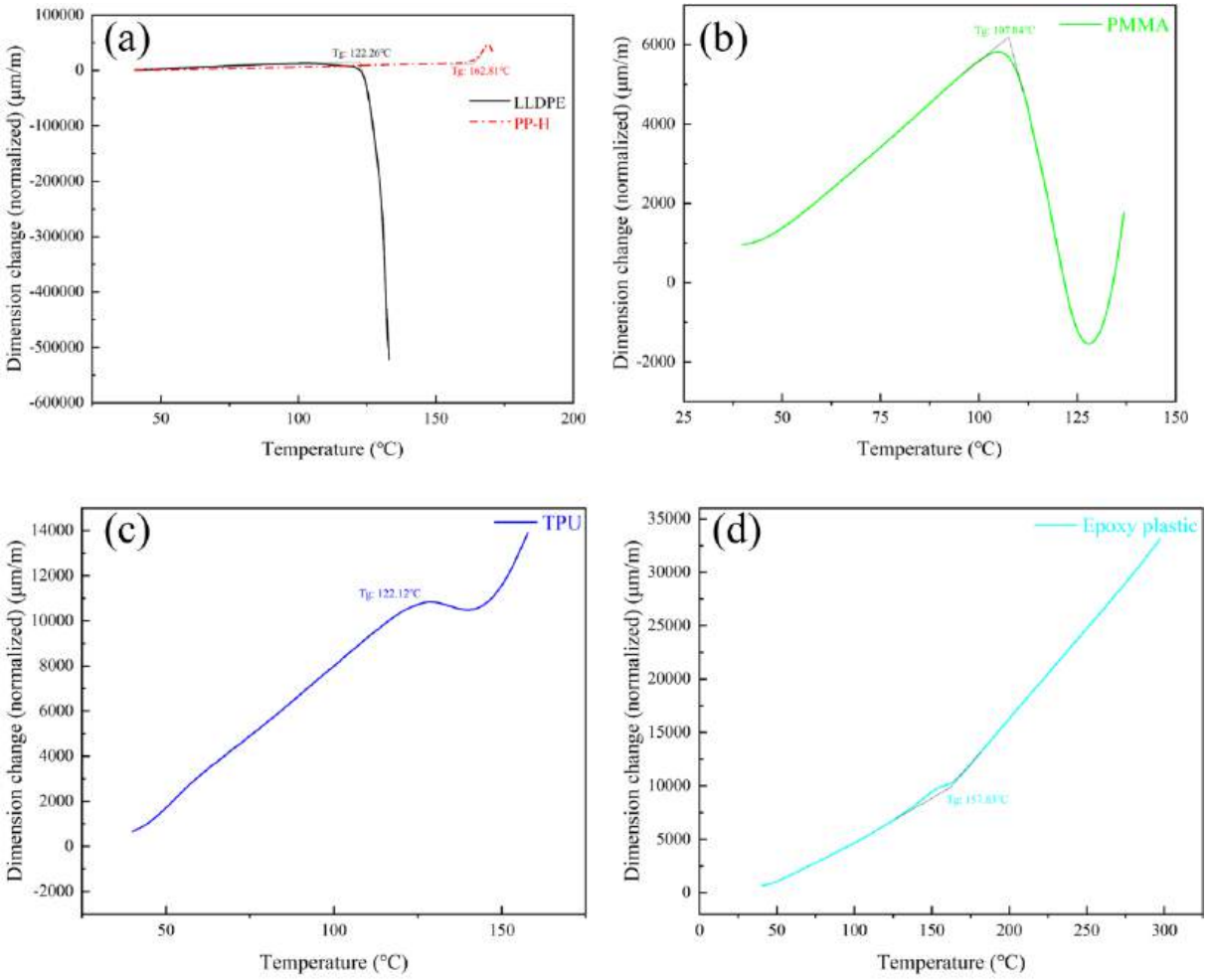


Fig. 3 Thermal mechanical analysis curves: (**a**) LLDPE and PP-H, (**b**) PMMA, (**c**) TPU, and (**d**) EP.

4. Conclusions

The mechanical properties (ultimate tensile strength, UTS; ultimate flexural strength, UFS; Izod notched impact strength, INIS) of LLDPE, PP-H, PMMA, TPU and EP polymers are tested in the temperature range from -60°C up to 80 °C. It was found that the mechanical properties of all materials are more or less temperature-dependent. The main conclusions are as follows:

1) The UTS of LLDPE, PP-H, PMMA and EP decreased with increasing temperature, possibly due to the enhancement of the movement ability of molecular chains. TPU was not fractured within the stroke of the testing machine, and its maximum tensile stress during stretching did not change much with temperature.
2) Only EP was fractured after the bending tests. Its UFS decreased first and then increased when the test temperature increased, reaching the minimum value of 102.4 MPa at 0℃.
3) As for the impact tests, ductile PP-H and TPU were not fractured, and the rest three materials were fractured after testing. As brittle materials, EP and PMMA exhibited

very low INISs at all testing temperatures. The INIS of LLDPE demonstrated a negative correlation to temperature.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

Kexin Deng and Derek W. Davies conceived the research idea. Jiaheng Li conducted the experimental investigation and data collection. Junhui Hu was responsible for project administration. Kexin Deng designed the methodology and performed the formal analysis. Jiaheng Li contributed to the review and editing of the manuscript. Derek W. Davies provided supervision throughout the project. All authors read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] T. Long, Q. Pang, Y. Deng, X. Pang, Y. Zhang, R. Yang, and C. Zhou, Recent progress of artificial intelligence application in polymer materials. Polymers **17**, 1667 (2025). https://doi.org/10.3390/polym17121667

[2] A. Sionkowska, Current research on the blends of natural and synthetic polymers as new biomaterials: Review. Progress in polymer science **36** (9), 1254 (2011). https://doi.org/10.1016/j.progpolymsci.2011.05.003

[3] G. Zhang, X. Cheng, Y. Wang, and W. Zhang, Supramolecular chiral polymeric aggregates: Construction and applications. Aggregate **4** (1), e262 (2023). https://doi.org/10.1002/agt2.262

[4] A. K. Nandi and D. P. Chatterjee, Hybrid polymer gels for energy applications. Journal of Materials Chemistry A **11** (24), 12593 (2023). https://doi.org/10.1039/d2ta09525d

[5] J. Pfleger and V. Cimrová, Polymers for electronics and photonics: Science for applications. Chemical Papers **72** (7), 1561 (2018). https://doi.org/10.1007/s11696-018-0472-0

[6] T. D. Huan, S. Boggs, G. Teyssedre, C. Laurent, M. Cakmak, S. Kumar, and R. Ramprasad, Advanced polymeric dielectrics for high energy density applications. Progress in Materials Science **83**, 236 (2016). https://doi.org/10.1016/j.pmatsci.2016.05.001

[7] J. Li, X. Liu, Y. Feng, and J. Yin, Recent progress in polymer/two-dimensional nanosheets composites with novel performances. Progress in polymer science **126**, 101505 (2022). https://doi.org/10.1016/j.progpolymsci.2022.101505

[8] R. Geyer, J. R. Jambeck, and K. L. Law, Production, use, and fate of all plastics ever made. Science Advances **3**, e1700782 (2017). https://doi.org/10.1126/sciadv.1700782

[9] L. Li, L. Zhong, K. Zhang, J. Gao, and M. Xu, Temperature dependence of mechanical, electrical properties and crystal structure of polyethylene blends for cable insulation. Materials **11** (10), 1922 (2018). https://doi.org/10.3390/ma11101922

[10] L. d. Reis and J. Marciano, Effect of temperature on the mechanical properties of polymer mortars. Materials Research **15** (4), 645 (2012). https://doi.org/10.1590/S1516-14392012005000091

[11] J. Richeton, S. Ahzi, K. S. Vecchio, F. C. Jiang, and R. R. Adharapurapu, Influence of temperature and strain rate on the mechanical behavior of three amorphous polymers: Characterization and modeling of the compressive yield stress. International Journal of Solids and Structures **43** (7/8), 2318 (2006). https://doi.org/10.1016/j.ijsolstr.2005.06.040

[12] Z. Z. Gao, Z. Q. Liu, W. Liu, and Z. F. Yue, Experimental and constitutive investigation on tensile and compressive behavior of MDYB-3 at different temperatures and loading rates. International Journal of Polymeric Materials **60** (5), 340 (2011). https://doi.org/10.1080/00914037.2010.531805

[13] W. Hu, H. Guo, Y. Chen, R. Xie, H. Jing, and P. He, Experimental investigation and modeling of the rate-dependent deformation behavior of PMMA at different temperatures. European Polymer Journal **85**, 313 (2016). https://doi.org/10.1016/j.eurpolymj.2016.10.036

[14] A. A. Abdel-Wahab, S. Ataya, and V. V. Silberschmidt, Temperature-dependent mechanical behaviour of PMMA: Experimental analysis and modelling. Polymer Testing **58**, 86 (2017). https://doi.org/10.1016/j.polymertesting.2016.12.016

[15] H. Ashrafizadeh, P. Mertiny, and A. McDonald, Evaluation of the effect of temperature on mechanical properties and wear resistance of polyurethane elastomers. Wear **368**, 26 (2016). https://doi.org/10.1016/j.wear.2016.08.008

[16] C. M. Tai, R. K. Y. Li, and C. N. Ng, Impact behaviour of polypropylene/polyethylene blends. Polymer Testing **19** (2), 143 (2000). https://doi.org/10.1016/S0142-9418(98)00080-4

Figure Captions

Fig. 1 Size diagrams of the mechanical property testing samples: (a) uniaxial tensile, (b) three-point bending, and (c) Izod notched impact.

Fig. 2 Temperature dependence of mechanical properties: (a) uniaxial tensile, (b) three-point bending, and (c) Izod notched impact.

Fig. 3 Thermal mechanical analysis curves: (a) LLDPE and PP-H, (b) PMMA, (c) TPU, and (d) EP.

Table Captions

Table 1 Ultimate tensile strength of the five investigated polymers at different test temperatures.

Table 2 Ultimate flexural strength of the five investigated polymers at different test temperatures.

Table 3 Izod notched impact strength of the five investigated polymers at different test temperatures.

From Plato's Cave to Algorithmic Feeds: Idealization and Moral Development in the Age of Social Media By Srihari Padmanabhan

Abstract

Since its emergence in the 2000s, social media has profoundly changed communication, social behavior, political engagement, and cultural norms. Yet, its historical evolution and the effect of algorithmic curation remain underexplored. This literature review examines how social media platforms, from early social networks like Classmates.com and Friendster to algorithmic programs like Instagram and TikTok, shape idealized perceptions of beauty, life, and success and influence moral reasoning. Drawing on Kohlberg's Theory of Moral Development and Plato's Theory of Forms, this review argues that social media platforms create feedback loops that arrest users at the conventional moral reasoning stage while amplifying different gendered experiences. Through historical analysis of social media evolution, review of psychological studies on social feedback and moral development, and philosophical interpretation on media idealization, our findings indicate that social media can function both as a moral and cultural force. This gives insight into moral development and critical thinking in contemporary society.

Introduction:

Social media has become a central part of our lives. Investopedia defines social media as "forms of digital communication through which users can share information and connect with online communities through texts, videos, or photos". Examples of prominent social media platforms include Facebook, LinkedIn, Twitter, Instagram, and TikTok. Recent studies have reported that social media can create "echo chambers" that reinforce existing beliefs (Haug et al) and quantify peer approval through likes, shares, and followers, which influence the user's decision–making (Flores & James).

Although previous studies have examined platform – specific effects on behavior and cognition, there is a lack of research toward the historical evolution of social media and the effect of algorithmic curation on moral development and gendered experiences. Prior studies have primarily focused on isolated social behaviors or a single demographic, giving rise to two critical questions. How do algorithms reinforce idealized standards of beauty, success, and social behavior? How does repeated exposure to idealized content affect moral development and gendered experiences over time?

In this literature review, we analyze the historical evolution of social media platforms, primarily focusing on the shift from early social networks in the 1990s to modern algorithmic platforms in the 2010s. We examine how these social media platforms curate content to increase user engagement and how they can influence moral development. Overall, our goal is to better understand the moral, ethical, and cultural consequences of social media while balancing the need for societal awareness and responsible platform design.

II. Historical Context of Social Media

The 1990s: Early Social Networks

After the development of the World Wide Web, several social networks were launched, which included Classmates.com in 1995 and SixDegrees.com in 1997. These were recognizable social networking sites, where individuals can create their profiles, list friends, and browse their social connections (Dhingra & Mudgal 3). Despite attracting several million users, this market was not fully prepared for mass social networking due to its restricted functions and limited possibilities (Castro et al. 36).

Between 1997 and 2001, several niche platforms emerged, which included Asian Avenue, Black Planet, and MiGente, that combined unique features, such as creating profiles, sending friend requests, and identifying friends on their personal profiles (Dhingra & Mudgal 3). In 1999, the Korean virtual site "Cyworld" and the Swedish web community "Lunarstorm" was launched, which offered features like guestbooks and friendlists.

The 2000s: User – Generated Content and Social Networking

Next Wave of Social Networks

With globalization, the popularization of the Internet, and the rise of the middle class in emerging countries, there was a greater demand for products that can facilitate relationships and social interaction. In 2001, Ryze.com was launched that helped users leverage their professional networks toward business (Dhingra & Mudgal 3). In 2002, Friendster was launched, "giving its users control over whom they wanted to connect rather than a computer–managed environment" (Dhingra & Mudgal 3). While Ryze never attained mass popularity and Friendster failed due to its technical and social problems, LinkedIn successfully distinguished itself by focusing on business services and helping people harness their social networks for jobs (Dhingra & Mudgal 3).

MySpace and the Customization Revolution

In 2003, MySpace gained popularity, leveraging on the rumor that Friendster would begin charging its users (Dhingra & Mudgal 4). It distinguished itself from other social networking sites by adding features and allowing users to personalize their pages. For instance, MySpace allowed users to customize their profiles, provide detailed information about their background and interest, and download music (Dhingra & Mudgal 4).

Facebook's Strategic Expansion

In 2004, Facebook emerged as a "Harvard - only site" and later expanded to universities, then to high school students, and then to those older than 13 years of age (Dhingra & Mudgal 3).

By 2006, Facebook became widely available to the users and was one of the popular social networking sites, alongside Pinterest, Tumbler, Spotify, etc.

Facebook's growth strategy proved to be quite successful. On October 24, 2007, Microsoft purchased 1.6% of Facebook share and gained rights to Facebook's international advertisements (Dhingra & Mudgal 4). By July 2010, Facebook reached 500 million users, and by the third quarter of 2017, this number increased to 2.07 billion active users (Dhingra & Mudgal 4). It gained popularity across Malaysia, Philippines, Australia, Singapore, Indonesia, Hong Kong, etc.

The 2010s: Algorithmic Acceleration and Hyper – Personalization

Twitter and Microblogging Revolution

Twitter, launched in 2006, but gained mass popularity in the 2010s, offered a possibility for users to follow each other and write small post messages with a maximum of 140 characters (Castro et al. 37). These features made Twitter the "go-to" for breaking news and opinions on current events, later becoming a platform for hate speech and tribalism.

Instagram and Visual – First Revolution

In 2010, Instagram was launched as a social media app for photo – sharing. By the first half of 2016, Instagram reached 500 million users (Dhingra & Mudgal 4). It has introduced several ad tools, formats, and a new interface. For instance, in March 2016, Instagram introduced its major interest–based algorithmic feed, shifting from posts in a chronological order to one that optimizes feeds based on interest and relationship with people. In late May and early June 2016, Instagram introduced the "Business Profiles" option, where it makes it easier to review content analytics and follower demographics and to create ads within the platform. In August 2016, Instagram introduced the "Instagram Stories" feature, where users can post videos or images that tell a story about their life or brand. They can add flare and excitement to their stories through text, GIF stickers, and video fillers.

III. Plato's Theory of Forms Applied to Media Idealization

A. Understanding Plato's Theory of Forms

Plato's Theory of Forms asserts that the physical world we perceive is a copy of a higher realm of perfect, eternal, and unchanging "Forms" (Matthias). These "Forms" exist outside of space and time and can only be accessible through our intellect and wisdom, not through our senses.

A famous example of Plato's Theory of Forms is the Allegory of the Cave. In this allegory, a group of people have been chained inside a cave since birth and can only see shadows of objects on the opposite cave wall. The real objects are held in the light behind them, but since they are chained, they are unable to turn around. Thus, these people begin to believe that these shadows are the only reality of the outside world. One day, one of the prisoners is freed and is

able to turn around and eventually leaves the cave. The prisoner realizes that these shadows were imperfect copies of the perfect ideas in the world of Forms (Matthias).

This allegory can give insight into media idealization in modern society. The shadows on the cave wall can represent filtered photos, sensational headlines, or targeted advertisements that project an illusion of reality that is not authentic. Just like the prisoners that mistake the shadows for actual reality, individuals repeatedly exposed to these idealized representations are more likely to believe that they accurately reflect life. The Cave represents social media platforms users use. Many social media platforms and algorithms influence what individuals are exposed to and how we perceive the world around us. Specifically, they reinforce idealized standards of beauty, success, and happiness. Over time, social media can create a distorted sense of reality, leading individuals to become disconnected from actual reality.

Furthermore, Plato emphasizes that ignorance is bliss. Even after seeing what happened to their fellow prisoner taken outside of the cave, they refused to leave the cave, as they preferred their accustomed roles and certainties to the temporary pain of the "blinding light of the sun", despite having a chance for a richer and a fuller life (Matthias). Similarly, individuals prefer this idealized content, as it is well – supported by the community, avoiding the risk of social isolation.

The freed prisoner represents someone who has gained knowledge of True Reality and the forms. In the context of social media, he is someone who questions sources and listens to alternative viewpoints, unlike the prisoners in the Cave. While this process can be uncomfortable, Plato argues that people should take a chance on this possibility to better understand the world around them.

B. Evolution of the "Form" Across Eras

The 1990s: Friends as Universal Platonic Ideal

Friends is a popular American sitcom, which aired on NBC from September 22, 1994, to May 6, 2004, lasting ten seasons with a total of 236 episodes. By the time it aired its final episode on May 6, 2004, Friends became a cultural icon, with 52.5 million viewers watching its series finale, making it the fifth most-watched series finale in television history (Estimated 51.1M). Beyond entertainment, Friends portrayed a new way of living life, presenting the idea that "all you need are good friends" and that you can construct families through choice to navigate the pressures of life and build lasting relationships (Jicha). Also, Friends had a global impact, with 26% of English learners considering it as the best show for learning their English (Fern). Through the lens of Plato's Theory of Forms, Friends can be interpreted as depicting a perfect "Form" of young adulthood, friendship, and urban life. It offers viewers a glimpse into a time of life, where comfort and companionship can be attained.

One of the most distinctive examples of idealization in Friends is Monica and Rachel's two–bedroom, one–bath apartment. Located in New York City's West Village, the apartment is

unusually spacious, brightly lit, and comfortably furnished, with an open kitchen/living room and an outdoor terrace. Given that Monica works as a chef at various restaurants with periods of unemployment and Rachel begins her career as a waitress before moving into the fashion industry, such a luxurious apartment is highly unrealistic. Yet, the show ignores the financial realities of urban life and presents the apartment as attainable and vital to the characters' lives. From the lens of Plato's Theory of Forms, the apartment can be seen as the "Form" of an ideal home. Its cozy atmosphere and vibrant neighborhood make it a place where comfort and security can thrive, rather than depicting actual reality.

Equally significant is the portrayal of the friend group. In Friends, these six characters function as a close – knit family, nearly spending all their free time together at Central Perk or their adjacent apartments. Despite their differing goals, professions, or even relationships, they value their friendship over their jobs or families, offering constant emotional support or help. Any disagreements or miscommunications between them are quickly resolved within an episode or two, and their lives are so intertwined that they experience all major life milestones (i.e., marriage, childbirth) together. But this cruel lie of "Friends", while emotionally appealing, gives unrealistic expectations of real friendships. In real life, modern adult friendships, especially in their 20s and 30s, are challenging to build and maintain. People begin to live elsewhere for new jobs or want to be closer to family. From a Platonic lens, the show's portrayal of the friend group can be seen as the "Form" of an ideal friendship, one that is always unconditional and permanent.

The show's six-character archetypes collectively form a Platonic "Form" of perfectly balanced social dynamics, where every interpersonal need is met, and every moral conflict is safely resolved. Rather than serving as complex, evolving individuals, they embody distilled personality traits. Chandler Bing, the Entertainer, utilizes sarcasm and wordplay to navigate complexity; Monica Geller, the Organizer/Hero, represents the relentless drive for structure and excellence; Ross Geller, the Geek, interprets mistakes through intellectual analysis; Rachel Green, the Connoisseur, encourages ambition and the pursuit of refinement; Joey Tribbani, the Neighbor, offers genuine, resilient warmth; and Phoebe Buffay, the Dreamer/Optimist, insists that creativity is always rewarded. Together, these character archetypes embody a social "Form", where no one is left out, reflecting Kohlberg's Stage 3 (Conventional Reasoning). Because the characters in *Friends* rarely face moral dilemmas that require them to choose abstract justice over group loyalty, the show creates a world where moral development is effectively arrested at the stage of "pleasing others." Repeated exposure to this idealized social Form suggests that happiness lies in the safe, static "Cave" of the community rather than the "blinding light" of independent, post conventional moral truth.

Repeated exposure to these idealized "Forms" of life can contribute to collective idealization, the consumption of unrealistic standards (i.e., lifestyle, appearance) across various digital platforms. However, this idealization is mediated through online communities, specifically forums and blogs. Both forums and blogs gave readers a space where they could ask

questions regarding the show's unrealistic economic realities or its polished version of urban life and form connections around these ideas. By allowing readers to like and follow these different ideas, viewers can appreciate these ideals and still be aware of their economic and social realities.

The 2000s: From Universal to Participatory Idealization

By the mid 2000s, the rise of social media platforms, such as MySpace and Facebook, has shifted media technologies from past anonymous usernames to personal profile–based identities. In the 1990 broadcast model, idealized versions of life were created by media institutions and passively consumed by users. However, in the mid 200s0s, users can produce and consume these idealized versions of life.

This shift greatly changed the role of Platonic "forms" in media culture. In the world of Friends, the six characters lived in improbably spacious apartments, maintained strong friendships, and navigated the pitfalls of work, life, and love with less long – term problems. In contrast, the real world contains isolation, loneliness, and even impossible problems, such as fear and menial work. Social media platforms help users escape financial and social constraints by allowing them to curate their lives themselves. Profile pages were carefully crafted to only highlight moments of happiness, success, and social belonging, while any signs of failure or insecurity were ignored. Thus, Platonic "forms" transformed from a singular universal ideal to various personal profiles with unique ideals.

Thanks to the rise of the "participatory" culture, it greatly affected the psychological relationship to idealization. Unlike the television characters in Friends, social media profiles belong to real people, making their idealized versions attainable and valuable. Users can now compare themselves against real people's curated versions of their lives, reducing the psychological gap between "Form" and reality.

The 2010s - Present: Algorithmic Amplification and the Invisible Hand

By the 2010s, the rise of smartphones and high – speed internet access has greatly reshaped the social media landscape. Initially, social media platforms would display posts in reverse chronological order to ensure users are updated with the latest content. As social media platforms expanded, the volume of content increased. As a result, the oversupply of content led to consumers decreasing their time spent online. In response, social media platforms utilize algorithms, aka the "invisible" hand, to curate content that matches the users' interests, enhancing the users' online experience.

This shift greatly changed the role of Platonic "forms" in media culture. In the 1990s, Friends presented a singular universal ideal of friendship and urban life. In the 2000s, curated profile pages contributed to the rise of a "participatory" culture, centered around sharing their curated versions with a targeted audience. By the 2010s, algorithms determine which ideals of success, beauty, or self – fulfillment is reinforced.

The Platonic form becomes increasingly personalized. Algorithms expose users to content that they find relevant or engaging while filtering out other posts, often without the users' conscious awareness. Over time, these idealized versions become social norms.

IV. Moral Development and Social Feedback Loops

A. Kohlberg's Theory of Moral Development Revisited

Kohlberg's theory (1984) is broken down into three primary levels: Pre– Conventional, Conventional, and Post– Conventional. Pre– conventional morality starts around infancy and lasts around age 9. At this age, the expectations of adults and the consequences of breaking the said rules shape the children's decisions. There are two stages within this level: Stage 1 (Obedience & Punishment) and Stage 2 (Individualism & Exchange). In Stage 1, young children must obey these fixed and absolute rules to avoid punishment. In Stage 2, young children must learn to account for individual perspectives and judge actions based on how they serve their human needs. Conventional morality starts around early adolescence and lasts till adulthood. There are two stages within this level: Stage 3 (Developing Good Interpersonal Relationships) and Stage 4 (Maintaining Social Order). In Stage 3, individuals live up to social expectations and roles to gain approval from others. In Stage 4, individuals attempt to maintain law and order by following rules or respecting authority. Post–conventional morality only occurs for some adults and focuses on individuals gaining an understanding of abstract principles of morality. There are two stages within this level: Stage 5 (Social Contract & Individual Rights) and Stage 6 (Universal Principles). In Stage 5, individuals develop abstract principles of justice to account for the differing opinions and beliefs of others. In Stage 6, people follow these abstract principles of justice, even if they conflict with preexisting laws and rules.

The Quantification of Peer Approval

Flores and James's research gives insight into the prevalence of approval – seeking behavior among youth in digital activities, such as social networks, blogging, content sharing sites, and online gaming. Drawing on qualitative interviews with 61 teens and young adults aged 15 - 25, their research found that "consequence thinking", a focus on gaining reward or avoiding negative consequences for the self, was the most dominant mode of reasoning, appearing around nine times per interview. Participants in this study described modifying their online behavior to avoid negative comments or to maintain a favorable image among their peers. For instance, athletes spoke about avoiding posting certain photos containing questionable content for fear of sanctions from coaches and school administrators (Flores & James 840).

A critical limitation of Flores and James's study is that its participants were aged 15 –25, a range for whom Stage 3 "consequence thinking" is developmentally typical, not necessarily arrested. However, Brady et al. addresses this gap by demonstrating that social learning processes can amplify moral outrage expression in a much broader adult population. Through an analysis of 12.7 million tweets from adult Twitter users, the researchers found that positive social

feedback (likes and shares) resulted in an expected 2% to 3% increase in the expression of moral outrage the following day (Brady et al. 3). This finding suggests that the digital "Form of the peer group can functionally freeze moral development at Kohlberg's Stage 3 by creating a continuous, algorithmically driven feedback loop that rewards social conformity over principled moral reasoning. Taken together, Flores and James and Brady et al. suggest that the social media's quantification of peer approval is not simply a developmental stage, but rather a tool that can sustain "consequence thinking" over time.

This pattern of "consequence thinking" is reflected in today's social media platforms. Whereas face–face interactions convey social approval through non-verbal cues, such as tone of voice or subtle nods, social media platforms use explicit, quantifiable metrics, such as likes, shares, or positive comments to achieve the same goal. Posts that receive 500 likes feel more validated and correct than posts that receive 50 likes, regardless of the content.

Furthermore, Flores and James's research gives insight into how these quantifiable metrics can shape users' behavior, even when the users have moral sensitivity. Their study observed that 75% of all consequence – thinking instances "pointed to the self," while the other 25% of interviewees pointed to other people (Flores & James 840). What this suggests for moral and ethical thinking is that peer feedback, now immediate and publicly visible, has encouraged users to anticipate how their content is received by their peers, rather than applying ethical principles. While users can recognize situations as moral in nature, they may not always act morally, as digital media can thwart the development of an autonomous self" and "destabilize relationships with others (Flores & James 836). Ultimately, participants recognize the potential harm their actions can do on others, but their moral sensitivity can be overridden by these quantifiable metrics found in social media. In turn, users may start adopting behaviors that others have adopted and perceive as beneficial to maximize their likes, shares, and positive comments, rather than critically evaluating the moral implications of their actions. Over time, this creates a feedback loop, where it amplifies conformity and prevents individuals from reaching post–conventional morality.

Flores and James's research found that moral thinking, a focus on the awareness of one's actions on others and adherence to principles of justice and equality, appeared in all participants at least once, with seven times per interview. Some expressed concern about how their actions can harm others, demonstrating person–centered moral thinking (40% of instances) (Flores & James 841). Others went further by taking the perspective of another person, demonstrating moral – perspective thinking (13% of instances). Still, others expressed a desire to treat all people with respect or see them in a moral way, demonstrating principled thinking (Flores & James 841). Yet, this "moral thinking" was restricted to the need to maintain relationships that benefit the "self" (Kohlberg's Stage 3), rather than embodying abstract thinking. Even participants who demonstrated principled thinking were unable to extend these principles to larger groups of people, which is necessary for ethical thinking (Flores & James 841).

Ethical thinking, the ability to think abstractly about the impact of one's actions on the larger community, was less prevalent and inconsistently used. 92% of participants exhibited community – level thinking at least once, averaging four instances per interview; 36% demonstrated roles and responsibilities thinking, yet "none of the young people interviewed appeared to engage ethical thinking consistently when making decisions online" (Flores & James 842). Quantified metrics in social media justify this inconsistency. Essentially, these metrics amplify content that is more well-liked than the others, regardless of the content. As a result, users will learn that having posts with critical evaluation reduces the likelihood of potential rewards or satisfaction.

Echo Chambers – Absence of Cognitive Conflict

Defining Echo Chambers and Filter Bubbles in Digital Contexts

Echo chambers are social networks where individuals can encounter opinions, beliefs, and information that amplify or reinforce their preexisting beliefs, while opposing or dissenting views can be minimized or ignored (Zabieno et al.109). Coined by Eli Pariser, the term "filter bubble", however, refers to the ways in which algorithms can tailor their search engine results and social media feeds ("intellectual isolation") to the users' choices and interactions (Talamanca & Arfini 2). Unlike "filter bubbles" that only operate in an algorithmic level, echo chambers can exist in offline settings, such as geographically isolated communities with similar views, but online platforms can increase membership numbers and have a greater significant impact (Talamanca & Arfini 7).

Both the "filter bubble" and the "echo chamber" can operate at different levels of abstraction. Pariser's "filter bubble" applies to consumption and exposure of online information at an individual level, while Sunstein's "echo chamber" applies to repeated social interactions within the online community (Talamanca & Arfini7). Both produce a feedback loop. Algorithms learn about user preferences and display "agreeable" content to increase the time the user spends on the platform (Rhodes 3). The "echo chambers" reinforce the homogeneous content by ensuring that the users interact with like–minded individuals (Zabieno et al 109). The users, repeatedly exposed to "agreeable" content, would be more likely to adopt these increasingly homogeneous views while ignoring dissent.

The Spiral of Silence and Echo Chamber Formation

Haug et.al's research gives insight into the psychological mechanisms driving "echo chamber" formation. Drawing on scenario – based survey that focused on three issues dominating the mainstream US landscape (immigration, presidential election, and COVID – 19), their study demonstrates that state – based fear of isolation (an emotional state arising from specific contexts) increases the likelihood that community members will promote the dominant opinion, irrespective of their views. These findings suggest that when users feel that their beliefs

don’t align with the dominant community’s opinion, they are less likely to express dissent and may end up promoting the prevailing opinion (Haug et al. 43).

This dynamic ties into the self – reinforcing nature of echo chambers. The fear of isolation pressures members to conform to community norms, and the algorithm, through the likes, shares, and positive comments, reinforces content that aligns with the members’ opinions. Thus, only certain opinions are accepted and amplified within the community, while others are either discredited through group criticism or ignored through lack of engagement (Haug et al. 33).

Moreover, the research found that “hot - button” issues (i.e., abortion, religion) are more likely to be “echo chambers” (Haug et al. 46). When the members are exposed to a dominant community of opinion that conflicts with their preexisting beliefs on highly salient topics, cognitive dissonance increases. The increase in cognitive dissonance increases state–based isolation, making members less likely to express dissent. Conversely, when a topic is less important, the members can freely express dissent, as they don’t have to worry about social isolation. Thus, high topic importance and fear of social isolation can allow for greater “spiral of silence” within communities.

Echo Chambers and Critical Thinking Deterioration

Research by Rhodes conducted a study to examine the effects of “echo chambers” and “filter bubbles” on the believability of political misinformation. Participants were randomly assigned to real and fake news articles, either Democrat or Republican – leaning, embedded within congruent or incongruent information streams.

The findings were concerning. When exposed to misinformation that is presented alongside “agreeable “material, both Democratic and Republican participants were more likely to rate fake news as believable, although it was slightly stronger for Republican participants. In contrast, Democratic participants were less likely to rate fake news as believable when exposed to incongruent information streams (Rhodes 11-12). Ultimately, the findings suggest that disrupting "echo chambers” can trigger systematic – based thinking for certain political groups (Rhodes 10). More importantly, the study found that exposure to “filter bubbles” can increase reliance on heuristic–based thinking (Rhodes 3). This finding suggests that “echo chambers” can affect the cognitive processes of individuals, ultimately reducing the users’ ability to critically evaluate political misinformation. This finding gives insight into moral development. Kohlberg suggests that only certain individuals who experience life–hanging experiences (i.e., ethical dilemmas, social movements) or think beyond societal rules can advance beyond conventional morality to post – conventional morality. However, “echo chambers” encourage individuals to fall back on familiar sources that confirm their preexisting views, leading to a lack of exposure to diverse perspectives and to critical evaluation of “questionable” material (Rhodes 5).

The cognitive consequences are significant. Firstly, users will start checking information based on whether it comes from "familiar" sources. Secondly, the users will struggle to synthesize opposing perspectives or allow for alternative viewpoints. Altogether, it prevents individuals from reaching post–conventional morality, where they consider how their choices can affect others and make decisions that benefit everyone.

VI. Counterarguments

A. Social Media Enables Post – Conventional Communities

A significant counterargument to this literature review is that social media does not simply arrest users at conventional moral reasoning. Instead, it can create post conventional communities that are centered around abstract principles of justice, equality, and universal rights. The most cited evidence for this claim is the rise of hashtag – based social movements, such as #MeToo and the # Black Lives Matter, which focus on using social media hashtags to promote social causes, raise awareness, and mobilize supporters around particular issues.

Empirical research by Mendes, Ringrose, and Keller examines how hashtag feminist activism can operate in practice. Drawing on an analysis of 800 pieces of digital content, such as blog posts or selfies, and qualitative interviews with 82 participants worldwide, which include organizers of feminist campaigns like #BeenRapedNeverReported and #MeToo, this case study documents how digital feminist hashtag activism can raise feminist consciousness and produce solidarity, corresponding with Kohlberg's post – conventional level of moral reasoning. Participants described coming to understand sexual violence not as an individual experience that arose from encounters with "bad men", but rather as part of a broader structural societal problem. For instance, one participant reported that sharing her experience online gave her the strength to report her assault to the campus police, showing how digital activism can provide opportunities for girls and women to conceptualize feminism as a personal and political imperative (Mendes et al. 238). Hence, the hashtag not only served as a platform for online expression but as a tool for channeling this collective awareness into concrete and meaningful social and political initiatives.

Furthermore, Mendes et al.'s empirical research notes that digital feminist platforms gave participants a political voice that many social contexts and structures, such as schools and workplaces, denied due to intersectional barriers related to race, class privilege, and age. 33% of survey respondents were teenagers attending school who reported that Twitter facilitated knowledge and dialogue that institutions could not (Mendes et al. 241). More importantly, 96% of the survey respondents reported directly challenging sexual violence and rape in their posts, suggesting an active engagement in feminist advocacy and social issues rather than submissiveness to preexistent social norms. However, Mendes et.al notes that this optimism carries essential qualifications. 72% of survey respondents reported experiencing negativity, hostility, or trolling in response to their feminist views, and many organizers reported attacks, burnout, and issues surrounding their work – life balance (Mendes et al. 242). These findings reveal while that post – conventional thinking can be possible through digital feminist

engagement, the platform design, which prioritizes virality and vitriol content, will amplify the preexisting social norms and ultimately ignore the participants' critical thinking that has allowed for social change.

B. Users Have Agency in Their Consumption

A second counterargument to this literature review is that users exercise agencies in how they engage with digital content. Rather than passively absorbing the platform norms, individuals actively interpret, select, and respond to the content they encounter in these social media platforms. According to the Uses and Gratifications Theory, individuals select and utilize certain types of media to satisfy specific wants and needs, such as acquiring information, experiencing aesthetic or emotional gratification, and strengthening relationships with family and friends (Katz, Blumler & Gurevitch). This perspective argues that active media users can control how much and in what ways media can influence their attitudes, beliefs, and behaviors.

Van Bevel et al. lends partial empirical support to this position. Their review acknowledges that users tend to express moral outrage for various motivations. Some are prosocial, such as raising awareness to a social injustice or mobilizing collective action, while others are self – serving, like pursuing social status through one's moral commitment, which is also known as "moral grandstanding" (Van Bevel et al 319). Essentially,"moral grandstanding" is a form of agency where individuals can curate others' moral discourse to boost their status and to make others believe that one is morally respectable. Furthermore, Van Bevel et al. notes that when users are encouraged to value accuracy over a partisan lens, their accuracy of people's beliefs and sharing decisions improves (329). This complicates the narrative that algorithms are the sole drivers for the spread of misinformation or the encouragement of outrage–driven behavior. While algorithms shape how users can encounter and share divisive and polarizing information, they do not entirely determine how users engage with this content. By designing algorithms and incentive structures in a way that can encourage the spread of productive content, users can be guided towards meaningful engagement, mitigating the impact of self–serving behaviors on their critical thinking.

Social Identity Theory provides an additional theoretical lens when understanding user agency. According to Social Identity Theory, individuals tend to identify into ingroups that positively contribute to their self–concept and compare their in–groups favorably against their outgroups (Tajel & Turner). In terms of social media, this means that users do not passively consume content but rather decide what type of content they want to participate in or the groups they want to surround themselves with.

Users can also exercise their agency through the ways they present their moral beliefs. Van Bevel et al. expresses that quantifiable social media metrics can serve as a type of reinforcement learning, where expressions of outrage can earn disproportionate positive social feedback, incentivizing users to amplify their outrage expressions over time (318). This suggests that users can make decisions on how they want to present themselves, even if they are exposed

to diverse, conflicting viewpoints from various out-groups. Moreover, Van Bevel et al. notes that a representative sample of Americans reported not wanting moral outrage and polarizing content to go viral on social media (322). This reveals a meaningful gap between the users' stated preferences and their digital behavior, suggesting that users can maintain their critical thinking on their engagement in the face of conspiratorial thinking.

VII. Synthesis and Implications

Through historical, psychological, and philosophical perspectives, we argue that social media platforms serve as active moral and cultural forces that influence how users perceive social reality and develop their moral reasoning.

Historically, the evolution of social media from early social networks in the 1990s to engagement – optimizing algorithms in the 2010s represents a shift in how users interact with idealized content. The 1990s broadcast model, exemplified by the popular American sitcom Friends, presented a singular, universal ideal of friendship, urban life, and success. Users would view this idealized content passively and critically, which is mediated through forums, blogs, and peer discussion. In the 2000s, users begin producing their curated self – portraits on MySpace and Facebook, reducing the psychological gap between "Form" and reality. By the 2010s, algorithms can now select and rank posts based on relevance or engagement without the user's conscious awareness or consent.

Philosophically, Plato's Theory of Forms and the Allegory of the Cave help us understand how algorithmic curation can distort users' perceptions of reality. Like the prisoners in the Allegory of the Cave, algorithmically curated users' mistake engagement – optimized content for social norms. The Forms reinforced by social media, such as idealized beauty, curated achievement, and meaningless friendships, are not random. Algorithms carefully select these Forms to increase attention and engagement among users, which means they are selected for their abilities to garner social approval and to follow social norms. As a result, the gap between representation and actual reality is maximized, while the user's conscious awareness of this gap is minimized. This does not mean that the user lacks cognitive ability but rather operates in an environment designed to prioritize engagement over accuracy.

Psychologically, Kohlberg's Theory of Moral Development provides a critical – thinking lens. Conventional moral reasoning, characterized by internalization of social norms and the tendency to judge right and wrong based on social approval or adherence to rules, represents a typical stage in moral development rather than a flaw. What social media platforms do, as this literature review argues, is integrating conventional moral reasoning into mainstream culture. The quantification of peer approval through likes, shares, and comments transforms Stage 3 consequence thinking from a developmental phase into a socially reinforced system that regulates users' actions and self – worth. As documented by Rhodes, Haug et al., and Zabieno et al., echo chambers and filter bubbles work towards erasing users' exposure to cognitive conflict and diverse perspectives, which Kohlberg identifies as essential for users to achieve post –

conventional moral reasoning. The result is simply an arrest, meaning that users who advance towards Stage 5 and Stage 6 abstract moral reasoning are trapped in an environment that rewards them for demonstrating conformity and punishes them for expressing dissent.

The counterarguments as seen in Section VI are genuine and empirically grounded. Mendes et al. notes real instances of post – conventional moral thinking emerging through hashtag activism, where participants begin to view sexual violence as a systemic injustice rather than a personal misfortune, paving the way for collective action. Van Bevel et al. expresses that users exercise agency in their social media consumption, primarily through expressing moral outrage for prosocial and self–interested causes and advocating for accuracy in social media posts. In addition, the Uses and Gratifications Theory (Katz, Blumler & Gurevitch) and the Social Identity Theory (Tajfel & Truner) add support to the idea that users actively engage with their social media communities.

Yet, it is important to note that the ideas of post–conventional communities and user agencies operate against social media platform design. Van Bevel et al. acknowledges a self–perpetuating outrage cycle in which algorithms encourage users to continue demonstrating their moral commitment irrespective of their fundamental values (320). Mendes et al. notes that 72% of feminist activists experienced hostility and trolling, and many reported burnouts from engaging in digital feminist activity (242). This suggests that behaviors, such as conformity, disengagement, and creation of new communities, are natural responses to an environment where they have little to no control over the content and feedback they encounter.

These findings give insight into practical implications for these three platform domains. At the platform design level, the literature review argues for algorithmic transparency and diversity programs that can expose users to heterogeneous content and disrupt the self–reinforcing cycle of the algorithms. Van Bevel et al. notes that shifting algorithmic focus towards accuracy rather than moral outrage can better improve the quality of information users encounter and share (329). At the educational level, having positive parental involvement and school–based support systems can help improve adolescents' media literacy and self–reflection, allowing them to navigate online risks responsibly and confidently. At the policy level, social media platforms should enforce responsible platform accountability to ensure that the platform design choices do not interfere with users' moral development.

VIII. Limitations and Future Research

This literature review acknowledges several limitations.

First, most of the studies rely on self–report and cross–sectional research, which makes it determine a cause-and-effect relationship and increases the risk of bias. Although Flores and James, Rhodes, and Haug et al. show connections between peer approval, echo chambers, and conventional moral reasoning, they do not fully explain this relationship. It is unclear whether excessive social media usage can arrest moral development, or if individuals at certain stages of moral reasoning are more drawn to social media. Future research should focus on using

longitudinal experimental designs, which track users' moral development over extended periods while manipulating exposure to social media algorithms, to yield accurate results.

Second, significant sampling limitations can affect the generalizability of many cited studies. Majority of published research on social media and moral reasoning has focused on Western, English – speaking, and affluent populations. Van Bevel et al. notes that 86% of research on social media and polarization is from the United States, even though nine out of the ten countries with the highest numbers of Facebook user populations are in the Global South. This geographical imbalance limits the extent to which these findings can accurately represent global patterns. Cultural context shapes the content users engage with, the norms reinforced within echo chambers, and the moral frameworks through which individuals evaluate peer approval, all of which can vary across various cultures, indicating that existing literature may not fully capture these differences. Future research should focus on sampling participants from non – Western and underrepresented communities to improve the generalizability of these studies.

Third, this review uses Kohlberg's Theory of Moral Development as its primary psychological framework, yet this theory, while influential, has its own limitations. Several critics have argued that Kohlberg overemphasized individual rights, abstract justice, and personal principles, which reflects the values of liberal, educated Western societies. However, in other cultural contexts, moral reasoning can prioritize harmony, tradition, or respect for authority, which Kohlberg's framework does not fully account for. This raises questions on whether his theory only considers cultures that fulfill Western liberalism and then project them as universal standards of morality. Additionally, Carol Gilligan, a prominent critic, has argued that the theory (1982) is gender – based, prioritizing a "morality of justice" (male – oriented) over "morality of care" (women – oriented), which emphasizes relationships, responsibility, and responsiveness to others' needs. This raises questions on which moral perspective is treated as the universal standard in psychological theories, and how moral and cultural voices can shape morality. This limitation becomes significant in this literature review, as social media platforms' quantification of peer approval can affect care – based and justice – based moral reasoning differently across social and cultural contexts in ways that Kohlberg's framework does not account for. Future research should incorporate Gilligan's Ethics of Care and Haidt's Moral Foundations Theory to better understand how algorithmic curation can shape users' moral psychology across multiple moral dimensions aside from justice.

Finally, this review has primarily focused on the moral developmental consequences of algorithmic curation for individual users. While institutional and political dimensions, such as platform shareholders and regulatory frameworks, are acknowledged, they are outside the scope of this review, thus limiting our ability to understand how systemic and structural factors can influence moral reasoning and online behavior. Future research should incorporate data from political economy or media studies to better understand how structural incentives of the global economy can shape platform design and to facilitate the creation of regulatory interventions that promote ethical algorithmic practices and support moral development at mass scale.

IX. Conclusion

This literature review examines how the historical evolution of social media platforms, from early social networks like Classmates.com and Friendster to algorithmic programs like Instagram and TikTok, has shaped users' moral reasoning and relationship with idealized content. Through Plato's Theory of Forms, we show that algorithms serve as an "Allegory of the Cave", which reinforce idealized perceptions of beauty and success that users mistake for reality. Through Kohlberg's Theory of Moral Development, we show that social media's quantification of peer approval through likes and shares rewards individuals with Stage 3 – consequence thinking, while echo chambers and filter bubbles remove critical – thinking, which is necessary for users to attain post – conventional moral reasoning. Even though social media has enabled post–conventional communities and users to exercise agencies in their social media platform, this functions within a platform that prioritizes engagement over moral reasoning. Our findings call for algorithmic transparency and responsible platform design to mitigate these effects.

Works Cited

Brady, William J., et al. "How Social Learning Amplifies Moral Outrage Expression in Online Social Networks." *Science Advances*, vol. 7, no. 33, 2021, https://doi.org/10.1126/sciadv.abe5641.

Castro, C., O. I. Reily, and A. Carthy. "The Evolution of the Internet and Social Media: A Literature Review." *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 12, no. 1, 2021, pp. 30-41, https://doi.org/10.17706/ijeeee.2022.12.1.30-41.

Dhingra, M., and R. K. Mudgal. "Historical Evolution of Social Media: An Overview." *International Conference on Advances in Engineering Science Management & Technology (ICAESMT)*, Uttaranchal University, 2019, http://dx.doi.org/10.2139/ssrn.3395665.

"Estimated 51.1M Tune in for Friends Finale." *Fox News Channel*, 7 May 2004. Archived 17 Dec. 2008.

Fern, C. "Students Learn English from 'Friends' on American TV." *Voice of America*, 30 July 2018. Archived 8 Aug. 2019.

Figà Talamanca, G., and S. Arfini. "Through the Newsfeed Glass: Rethinking Filter Bubbles and Echo Chambers." *Philosophy & Technology*, vol. 35, no. 20, 2022, https://doi.org/10.1007/s13347-021-00494-z.

Flores, A., and C. James. "Morality and Ethics Behind the Screen: Young People's Perspectives on Digital Life." *New Media & Society*, vol. 15, no. 6, 2013, pp. 834-852, https://doi.org/10.1177/1461444812462842.

Gilligan, Carol. *In a Different Voice: Psychological Theory and Women's Development*. Harvard University Press, 1982.

Haug, M., et al. "Supporting Opinions to Fit In: A Spiral of Silence-Theoretic Explanation for Establishing Echo Chambers and Filter Bubbles on Social Media." *Internet Research*, vol. 35, no. 7, 2025, pp. 30-51, https://doi.org/10.1108/INTR-03-2024-0413.

Jicha, T. "They Leave as They Began: With a Buzz." *The Baltimore Sun*, 2 May 2004, p. 1. Archived 4 June 2011.

Katz, Elihu, Jay G. Blumler, and Michael Gurevitch. "Utilization of Mass Communication by the Individual." *The Uses of Mass Communications: Current Perspectives on Gratifications Research*, edited by Jay G. Blumler and Elihu Katz, Sage Publications, 1974, pp. 19-32.

Kohlberg, Lawrence. *The Psychology of Moral Development: The Nature and Validity of Moral Stages*. Harper & Row, 1984.

Lauer, M. "'Friends' Creators Share Show's Beginnings." *Dateline NBC*, 5 May 2004. Archived 2 Oct. 2013.

Matthias, Andreas. "Plato's Theory of Forms: Stepping Out of the Shadows and into the Light." *Daily Philosophy*, 13 Dec. 2023, https://daily-philosophy.com/platos-theory-of-forms/.

Mendes, K., J. Ringrose, and J. Keller. "#MeToo and the Promise and Pitfalls of Challenging Rape Culture Through Digital Feminist Activism." *European Journal of Women's Studies*, vol. 25, no. 2, 2018, pp. 236-246, https://doi.org/10.1177/1350506818765318.

Tajfel, Henri, and John C. Turner. "An Integrative Theory of Intergroup Conflict." *The Social Psychology of Intergroup Relations*, edited by William G. Austin and Stephen Worchel, Brooks/Cole, 1979, pp. 33-47.

Rhodes, S. C. "Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation." *Political Communication*, vol. 39, no. 1, 2022, pp. 1–22, https://doi.org/10.1080/10584609.2021.1910887.

Van Bavel, Jay J., et al. "Social Media and Morality." *Annual Review of Psychology*, vol. 75, 2024, pp. 311-40, https://doi.org/10.1146/annurev-psych-022123-110258.

Zabieno, A. S., Damayanti, D., & Abdullah, A. Z. (2025). The Role of AI, Filter Bubbles, and Echo Chambers in Political and Religious Polarization on Social Media. *Dinamika Penelitian: Media Komunikasi Penelitian Sosial Keagamaan*, *25*(2), 102–118. https://doi.org/10.21274/dinamika.2025.25.2.102-118

A multi-criteria evaluation of nuclear reactor technologies for Kazakhstan's first nuclear power plant.

By Alikhan Zhetpis and Adel Abzhan
Zhetpis Alikhan, NIS of Astana, alikhan.zhetpis@gmail.com

Acknowledgements
The author would like to thank Dick Ferieno Firdaus from University of Cambridge for his invaluable contribution to this research paper.

Abstract - **Kazakhstan plans to construct its first nuclear power plant in the near future. This decision is motivated by energy sustainability concerns and Kazakhstan's 2060 carbon neutrality goal, which aims to reduce carbon emissions from power generation. The nuclear power plant initiative was broadly supported by the population of Kazakhstan in the Referendum held in October 2024. However, nuclear power plant construction raises significant concerns related to safety, security, water availability, financial risks and environmental contamination. The choice of the reactor technology is therefore crucial, as it determines the scale of these potential risks. This research evaluates the suitability of selected nuclear reactor technologies for Kazakhstan's first nuclear power plant using a multi-criteria assessment. The study compares Rosatom's VVER-1200, China National Nuclear Corporation's HPR-1000, French EPR-1200, and GE Hitachi's BWRX-300 across six criteria: safety and security, climate and water resources, grid compatibility, economic costs, fuel cycle and waste management, and institutional fit. The results show that HPR1000 receives the highest weighted score of 3.80/5.00, indicating the most balanced overall suitability. BWRX-300 performs strongly in grid and water compatibility but remains commercially uncertain. VVER-1200 is practical but politically dependent, while EPR1200 faces major cost and delivery risks. Overall, the findings demonstrate the importance of a multi-dimensional approach in the selection of reactor technology, highlighting the significance of technical, ecological, economic and institutional factors.**

Key terms: Energy sustainability; grid compatibility; Kazakhstan's first nuclear power plant; water resources.

1. Introduction

Climate change has intensified significantly over the past years, becoming one of the most pressing global issues of the 21st century. According to Rohde, the last 11 eleven years, 2015-2025, have been reported as the warmest 11 years ever to have been directly observed, which validates the growing public concern. Climate change is driven by the excess production of greenhouse gases such as primary greenhouse gas CO_2, CH_4, N_2O. A large contribution to this excess production can be attributed to non-renewable sources of energy, which include coal, oil and natural gas: contributing 41%, 32% and 21% of global CO2 emissions respectively (Ritchie et al.). The escalating threat posed by climate change has prompted coordinated measures from the international community. This resulted in the Paris Agreement, signed by 195 countries on December 12, 2015. It aims to decelerate climate change by reducing greenhouse gas emissions; by keeping global temperature rise below 2°C and striving for 1.5°C; aligning financial flows with low-carbon development (Burck et al.).

Kazakhstan, as a signatory to the Paris Agreement, targets the achievement of carbon neutrality by 2060 (President of the Republic of Kazakhstan). In 2024, Kazakhstan demonstrated substantial reliance on fossil fuels for energy extraction: 51.5%, 29.9% and 17.7% of total primary energy production came from oil, coal and gas sources respectively; while only 6.2% was generated using renewable sources of energy ("Fuel and Energy Balance"). This illustrates the need to transition to alternative sources of energy. Means of implementing carbon neutrality necessitate the development of zero-carbon emission energy sources. One of the most promising energy sources is nuclear energy. Staggering efficiency and independence of weather conditions distinguishes it from other sources of carbon-clean energy. Consequently, development and construction of nuclear powerhouses is viewed as a logical pathway towards achieving the 2060 objective among Kazakhstan's society.

Accordingly, the national referendum about the construction of a nuclear power plant was conducted in October 2024. 71.1% of participants supported the building of a nuclear power plant and as such, the future construction of the facility was announced (Satubaldina). Thus, the process was initiated. Despite all the advantages presented by the nuclear power plant, its construction is associated with many challenges. This includes high economical costs, environmental considerations and energy grid compatibility (Batyrbekov et al.). Choosing an appropriate nuclear reactor type directly affects the aforementioned factors as the technology and design are unique and individual for each reactor. This emphasizes the importance of the considerate choice regarding the appropriate reactor type for the future power plant.

Several researchers have already examined the possibilities of implementing the nuclear program in Kazakhstan. Kim evaluated the importance of different factors contributing to the choice of nuclear reactor technology. Factors were categorized into the following groups: socio-political, technical, economic and environmental. The methodology included a survey among Ministries' representatives and workers. Although Kim found that environmental and economic factors carried the greatest relative weight, this conclusion is limited as the study analyzed only the governmental perspective.

Earlier reactor-selection studies provide a useful foundation for choosing an appropriate reactor type for Kazakhstan. For instance, Pospelov et al. conducted a multi-criteria evaluation to determine the most suitable technology. The study implemented detailed, highly technical and extensive criteria for a comparison of the reactors. Researchers arrived at the conclusion that Russian Generation III+ reactors VVER-1200 and VVER-1000 were the best fit for the future nuclear program. However, Pospelov et al. performed the research before the 2024 referendum and prior to the rapid development of small modular reactors, which presents a major drawback of the study. More recent academic work by Matebay addresses the aforementioned constraints by providing substantial analysis of miscellaneous reactor technologies, incorporating both multi-criteria evaluation and case study. To a large extent, this approach proved successful. Matebay compared EDF, CNNC, KHNP, Rosatom and Westinghouse companies across a wide range of characteristics, including technical, geopolitical and economical aspects. The study recommended Chinese Hualong Pressurized Reactor (HPR) reactors for Kazakhstan's nuclear power plant construction due to their extensive expertise in reactor construction, rapid project execution, competitive financing options, and ability to support the nuclear fuel cycle comprehensively (Matebay). Despite its in-depth analysis, the study excluded Small Modular Reactors (SMRs) from the evaluation, which represents the limitation of Matebay's study.

Recent research regarding SMRs explores the prospects offered by this innovative technology. Insepov et al. investigated the significance of SMRs in the context of future global nuclear development. Important advantages such as adaptability, passive safety and modular construction are limited by the economic competitiveness and lack of reference experience. The study performed by Insepov et al. hasn't provided any Kazakhstan-specific analysis of future SMR development, which is a substantial disadvantage of this paper when considered within this context.

In general, the existing literature either presents abstract analysis of nuclear reactor technologies without fully integrating Kazakhstan's conditions or provides insufficient, outdated evaluation of the nuclear reactor types. Thus, there remains a research gap in determining the most appropriate nuclear reactor technology for Kazakhstan's future construction of its first nuclear power plant. This research aims to examine this gap by implementing multi-criteria evaluation and literature review. Inspired by existing literature, the following criteria were developed for the evaluation: safety and security, climate and water resources, grid compatibility, economic costs, fuel cycle, waste management and institutional fit. Recognizing that each criterion contributes differently to the selection of reactor technology, weightings are introduced in accordance with the relative significance of each criterion. The present research poses a particular relevance since Kazakhstan's nuclear program is at the early stages of development, emphasizing the need for proper and nuanced guidance.

2. Literature Review

2.1 Working Principle of Nuclear Reactors

A nuclear power plant generates electricity with steam turbines, which utilize the thermal energy released during atomic fission inside the core of the reactor. Although the operating principle is common for all types of nuclear reactors, different designs implement various fuel, coolant, moderator, and other technical choices in the reactor. This leads to a great variety in nuclear reactor designs. Water-cooled - pressurized water reactor (PWR), boiling water reactor (BWR); gas-cooled - advanced gas-cooled reactor (AGR); with no moderator - fast neutron reactor (FNR) ("Nuclear Power Reactors").

The most widespread type of nuclear reactor design is PWR, which is used for almost 70% of the global reactor fleet ("Nuclear Power Reactors"). PWRs implement light water (H_2O) as both coolant and moderator, allowing for greater efficiency of thermal transfer and, relative to other reactor designs, cheaper construction cost. Additionally, the reliability and the inherent safety of PWRs largely contribute to its popularity in the nuclear market.

The basic structure and energy-conversion process of a pressurized water reactor are shown in Figure 1

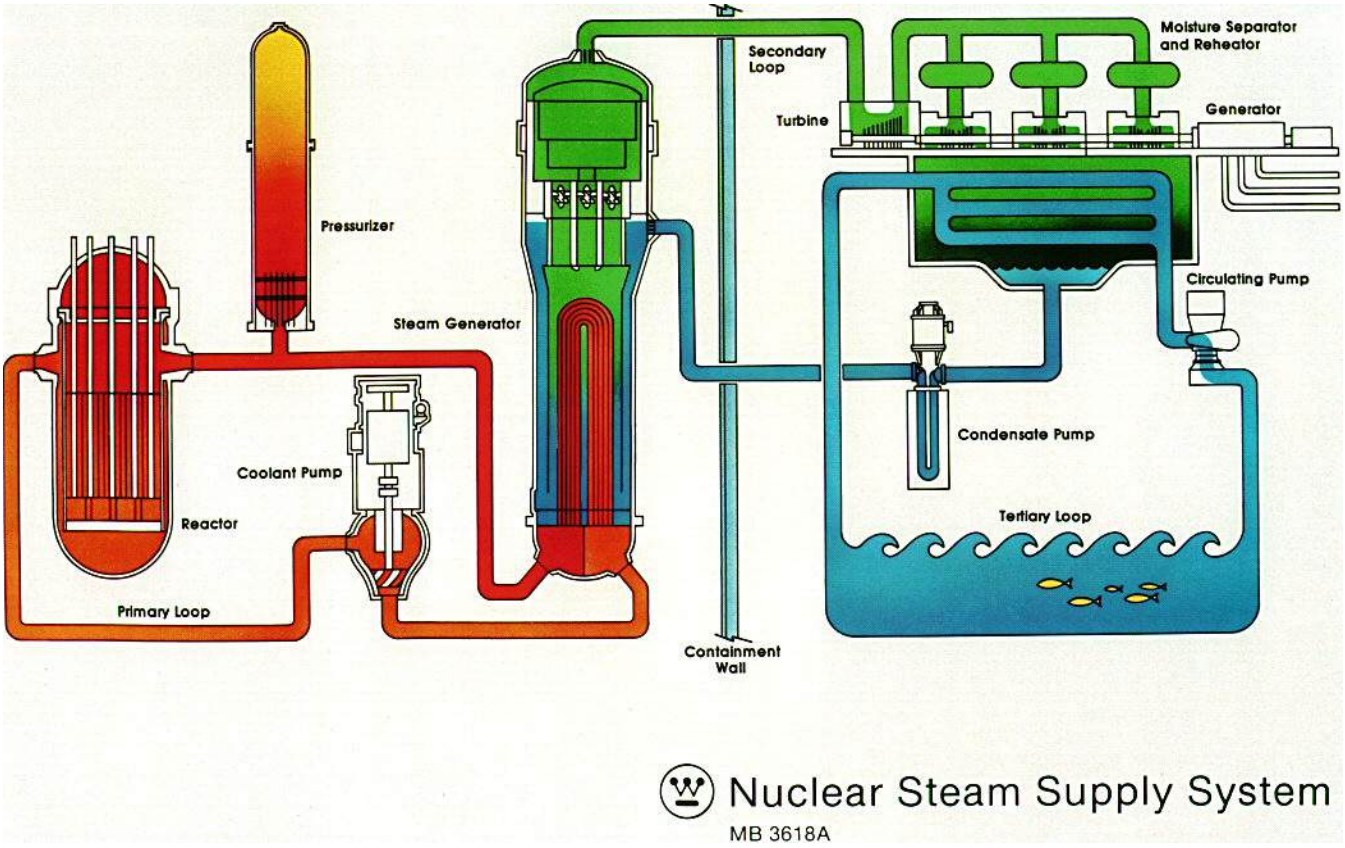


Figure 1. *Diagram of PWR Components.* This diagram shows the main components of a pressurized water reactor, including the reactor vessel, steam generator, pressurizer, coolant pump, turbine, condenser, and cooling system. Source: Joseph Gonyeau, *PWR Cycle*.

The operating principle of PWRs shares many similarities with other nuclear reactor designs. PWR employs the process of continuous nuclear fission to generate a substantial amount of heat energy required to power steam generators and produce

electricity. Fundamentally, nuclear fission is a process involving enriched uranium-235 and the slow-paced neutron. When a neutron hits uranium-235, it merges with the atom, creating unstable uranium-236. Uranium-236 breaks down into fission products, such as xenon, iodine, cesium, emitting 2-3 neutrons and energy. Released neutrons may induce fission reaction in nearby atoms of uranium 235. This causes the continuous chain reaction which drives the nuclear reactor ("Nuclear Power Reactors"). This reaction is illustrated in Figure 2.

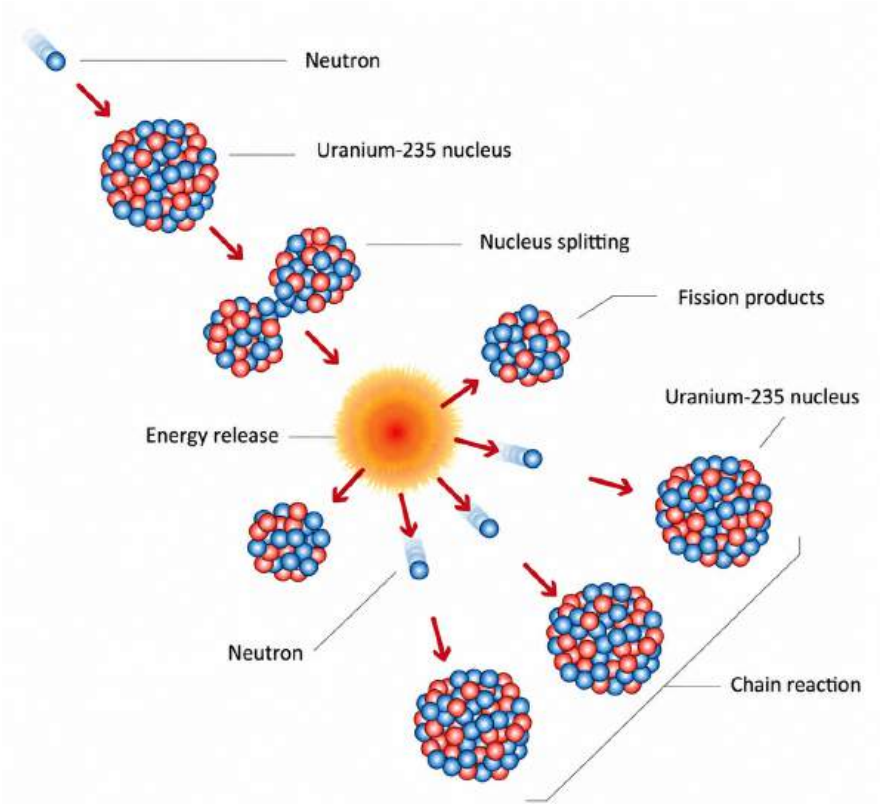


Figure 2. *Nuclear Fission Chain Reaction.* This diagram illustrates how a neutron initiates the fission of a uranium-235 nucleus, releasing energy, fission products, and additional neutrons that continue the chain reaction. Source: *Nuclear Fission Chain Reaction*, Nuclear Power.

The energy released during fission reaction is utilized using several light water circuits as shown in Figure 1. Light water (H_2O) is a great moderator for the nuclear reactor, as it is able to slow down the neutrons effectively, maximizing fission reaction probability. Uranium-235 fission occurs more easily with thermal (slow) neutrons compared to faster neutrons, as shown in Figure 3. Without moderators, fast-speed neutrons will travel too quickly for the interaction with fuel nuclei, severely decreasing the efficiency of the reactor. Another reason for the usage of light water is a high specific heat capacity, which is essential for thermal transfer ("Physics of Uranium and Nuclear Energy").

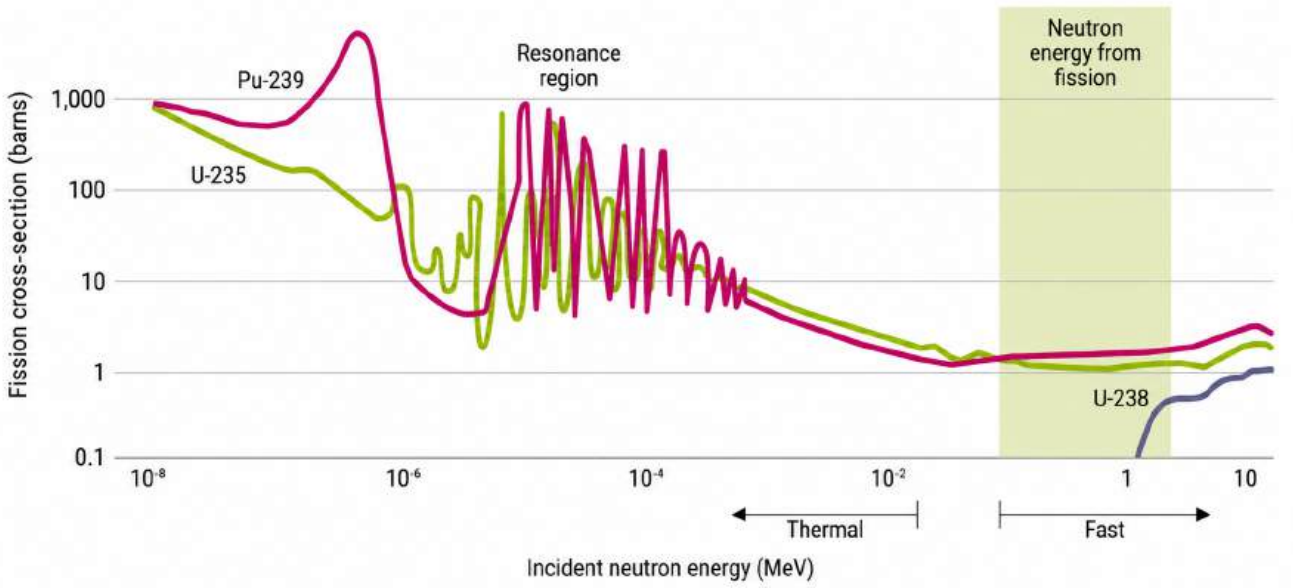


Figure 3. *Neutron Cross-Sections for Fission of Uranium and Plutonium.* This graph compares the fission cross-sections of uranium and plutonium isotopes, showing how their likelihood of undergoing fission varies with neutron energy. Source: World Nuclear Association, *Physics of Uranium and Nuclear Energy*.

PWRs usually implement two or more cooling circuits. This ensures safety and reliability of the reactor, and prevents radiation leakages. Using geometrical characteristics of rod and pellets, the rate of the fission reaction is sustained at such a level that it remains controllable. The controlled balance between increasing and decreasing the rate of fission reaction is key to power generation in PWRs ("Nuclear Power Reactors").

2.2 Overview of Nuclear Reactor Technologies

Generation III+ nuclear reactors, while sharing a common light-water moderated fission principle, differ significantly in engineering design, which directly affects the suitability of each individual reactor for Kazakhstan. One of the most prominent PWR designs is the VVER-1200 (Russia). It is a four-loop pressurized water reactor producing around 1,200 MW with 37% thermal efficiency and a 60-year design life, distinguished by horizontal steam generators. VVER-1200 utilizes a Passive Heat Removal System (PHRS) capable of maintaining decay heat removal for approximately 72 hours without external power (Rosatom, "VVER-1200"; "Nuclear Power Reactors"). It uses low-enriched UO_2 fuel (4–5% U-235) in a conventional open fuel cycle, requiring long-term spent fuel management. Its capital cost is relatively competitive, especially with Russian state financing (Markopoulos and Karkazis).

The EPR-1200 (France) represents a highly conservative Gen III+ design prioritizing maximum redundancy and extremely low core damage probability (10^{-7} per reactor-year) through a four-train active safety system, double containment, and ex-vessel corium spreading and cooling (EDF). With 1,200 MW output and 37% efficiency, it supports flexible fuel cycles, including mixed oxide fuel (MOX), but is characterized by high construction complexity and cost, as evidenced by past European projects. While offering one of the highest safety margins, its economic risk, construction delays, and operational complexity reduce its attractiveness for a first-time nuclear country. Waste management challenges remain similar to other Light Water Reactors (LWRs), requiring long-term storage or reprocessing strategies (Matebay).

The HPR-1000 (China) is a three-loop PWR producing 1,000–1,100 MW with 36% efficiency, combining active and passive safety systems, including passive injection, containment cooling, and a core catcher. It features a longer 18–24 month refueling cycle, improving operational efficiency, and offers competitive capital costs, particularly with Chinese state financing. Its slightly smaller unit size improves grid compatibility compared to larger reactors, and it provides geopolitical

diversification. However, it has less long-term international operating experience and more limited regulatory acceptance in Western markets. Like other LWR, it operates on an open fuel cycle, requiring long-term spent fuel management (Xing et al.).

In contrast to large reactors, SMRs - including designs such as BWRX-300 - offer outputs between 50–300 MW and are characterized by modular construction, integral primary systems, enhanced passive safety, and potential for air or hybrid cooling (International Atomic Energy Agency). These features make them particularly suitable for Kazakhstan's water-constrained and regionally fragmented grid, as well as for replacing aging coal plants. However, SMRs face uncertainties in first-of-a-kind costs, licensing, and fuel supply. Overall, reactor technologies differ not only in size but in cooling requirements, safety philosophy, economic viability, and fuel-cycle implications, making technology selection highly context-dependent rather than purely generational.

2.3 Environmental Conditions of Kazakhstan

After the national referendum in 2024 was held, the Kazakhstan government proposed and later approved the village Ulken, situated near the coast of Lake Balkhash, as an official site for the future construction of the first nuclear power plant. Siting of the nuclear facility has a great effect on the operating process of the reactor due to the importance of stable water supply. Also, regarding safety and security concerns, extreme climate and seismic conditions have to be considered when selecting an appropriate reactor technology.

Lake Balkhash is a closed lake located in southeastern Kazakhstan. It is characterized by its unique dual water composition: one part of the lake, fed by the Ile river, consists of fresh water, while the other part remains saline. Lake Balkhash has experienced substantial water level fluctuations over recent years and future projections indicate increasing water-resource stress, which is important for assessing the proposed nuclear power plant near Ulken. Alimkulov et al. project that by 2050 surface inflow may decline by about 30%, while evaporation may increase by about 25% compared with the 1981–2010 climate norm. Yang et al. provide a longer-term warning: even while assuming the most positive energy development scenario, the lake may decline by about 2.5 m from its 2020 level by 2100, while under worse scenarios it may approach its historical low level of 340.52 m by the 2070s and 2050s respectively. The main drivers are rising evaporation, unstable glacier and snowmelt contribution, irrigation demand, reservoir regulation, and transboundary water abstraction from the Ili River basin. Yang et al. describe the lake's recent stability as a "masked vulnerability," where climate-driven short-term gains have been negated by irrational human water use. Therefore, a nuclear power plant near Ulken could increase environmental pressure. However, the risk would be reduced if water-efficient cooling technologies or smaller capacity are implemented in the nuclear reactor.

Lake Balkhash is also susceptible to the seismic activity. Tengrinews.kz reports that Lake Balkhash's shallow depth causes only moderate seismicity, but Sokolova et al. show that the region is not fully aseismic, citing a body-wave magnitude (mb) = 4.5 earthquake in 1988 and a regional P-wave magnitude scale (mpv) = 4.8 tectonic earthquake in 2003 near Kounrad, felt in Balkhash at 3–4 intensity points. 3–4 points refers to weak-to-noticeable felt shaking. Thus, a nuclear reactor built at Ulken site has to account for the potential seismic activity to ensure the security of the plant.

2.4 Current Energy Needs

Compatibility of the future nuclear power plant with the energy sector of Kazakhstan substantially contributes to the choice of the appropriate reactor technology. Capacity and connectivity of the energy grid constitute the required power generation of the plant in order to prevent system overload. This factor appears of particular importance when selecting between modular reactor design or conventional large-scale reactor design.

As of 2025, Kazakhstan's energy sector is characterized by a growing imbalance between the consumption and production of electricity. According to Kazakhstan electricity grid operating company (KEGOC), the electricity generation of 123,110.9 million kWh was outmatched by 124,606.5 million kWh electricity consumption. This difference was covered by importing the electricity from Russia. Although relatively insignificant, amounting for 1.2% of total electricity consumption, this

discrepancy demonstrates the emergence of the structural supply gap in the national energy system. Combined with the deterioration of the local thermal power plants, whose overall level of wear and tear reaches 59% across the country, this threatens the energy security of Kazakhstan ("Energy Storage Systems").

Regarding the sustainability of the energy sector, Kazakhstan is far from its 2060 carbon neutrality goal. The share of electricity generated from renewable sources of energy was 6.2% of the total energy production in 2024 ("Fuel and Energy Balance"). Compared with 0.6% in 2015, the year the Paris Agreement was signed, this is a notable increase in the share of renewables within the national energy sector. This illustrates the gradual shift towards alternative, emissions-free sources of energy. It is further reflected in the governmental plan of development of the energy grid. By year 2035, Kazakhstan aims to incorporate 26 GW of new generating capacity through the construction of additional energy facilities; 3.8 GW will be supplied by renewable sources of energy ("Phased Introduction of New Capacities"). Generally, the share of electricity produced through renewable energy sources is expected to reach 15% in 2030. Nonetheless, the development of renewable energy alone may be insufficient for achieving the 2060 carbon neutrality goal, indicating the need for other, reliable and effective low-carbon alternatives.

The need to install additional generating capacity is further supported when Kazakhstan's inter-country power flows are reviewed. In 2025, Kazakhstan both imported and exported electricity to its neighboring countries: net imports from Russia amounted to 2,479.3 million kWh, while net exports to Central Asia totaled 983.7 million kWh (KEGOC). This presents both challenges and opportunities for Kazakhstan's energy sector. Reliance on electricity imports demonstrates the country's vulnerability in terms of energy security. This is evident during annual peak loads, when 810 MW, approximately 4.6% of total maximum hourly consumption, was supplied through imports (KEGOC). At the same time, exporting electricity provides the government with a supplementary source of income, potentially generating substantial revenue. In the long term, nuclear capacity could reduce import reliance and create export opportunities for Kazakhstan, depending on the grid load, energy demand and electricity pricing.

The uneven distribution of energy production over the country also strengthens the argument for new stable generating capacity. Kazakhstan's Unified Electric Power System (UES) is divided into three zones: Northern, Southern and Western. Northern and Southern zones are closely interlinked together, while Western zone is independent from the other two, supplied by its own plants and connected to Russian power flow. In 2025, electricity consumption was concentrated mainly in the Northern Zone, which accounted for 75,540.6 million kWh, or 60.6% of national consumption, while the Southern and Western Zones consumed 30,641.0 million kWh and 18,424.9 million kWh respectively (KEGOC). Combined with Figure 4, which shows that the Northern Zone also dominated the production of the electricity, amounting for 86.5 billion kWh or 73.4% of the total generated energy, this shows the heavy reliance of zones on covering regional deficits through transmission from other zones. It is especially relevant for the South, which covered 14.1 billion kWh of deficit from the Northern zone. In this context, a potential nuclear power plant could help reduce the Southern Zone's supply deficit by providing reliable baseload generation near a major demand region.

Although Kazakhstan currently lacks energy storage systems, this does not necessarily prevent the integration of nuclear power, since balancing can also be supported through grid connectivity and power transmission. KEGOC's planned projects are therefore significant: the South Kazakhstan Electricity Transmission Reinforcement Project is expected to increase North–South transit capacity by 440 MW, while the future North–South high-voltage direct current (HVDC) line is planned to expand transmission capacity by 2,000 MW (ENERGY Insight & Analytics; United Nations Economic Commission for Europe). Moreover, the future integration of the Western Zone into the unified power system by 2040 is proposed (United Nations Economic Commission for Europe). At the same time, energy facilities in Kazakhstan lack energy storage systems; previously, such systems were regarded as too expensive and were insufficiently incorporated into energy infrastructure planning. According to the Ministry of Energy, Kazakhstan currently has no operating energy storage systems, although future large projects are expected to include storage equal to 30% of station capacity; potentially, storages can be used for nuclear facilities as well due to their compatibility ("Energy Storage Systems").

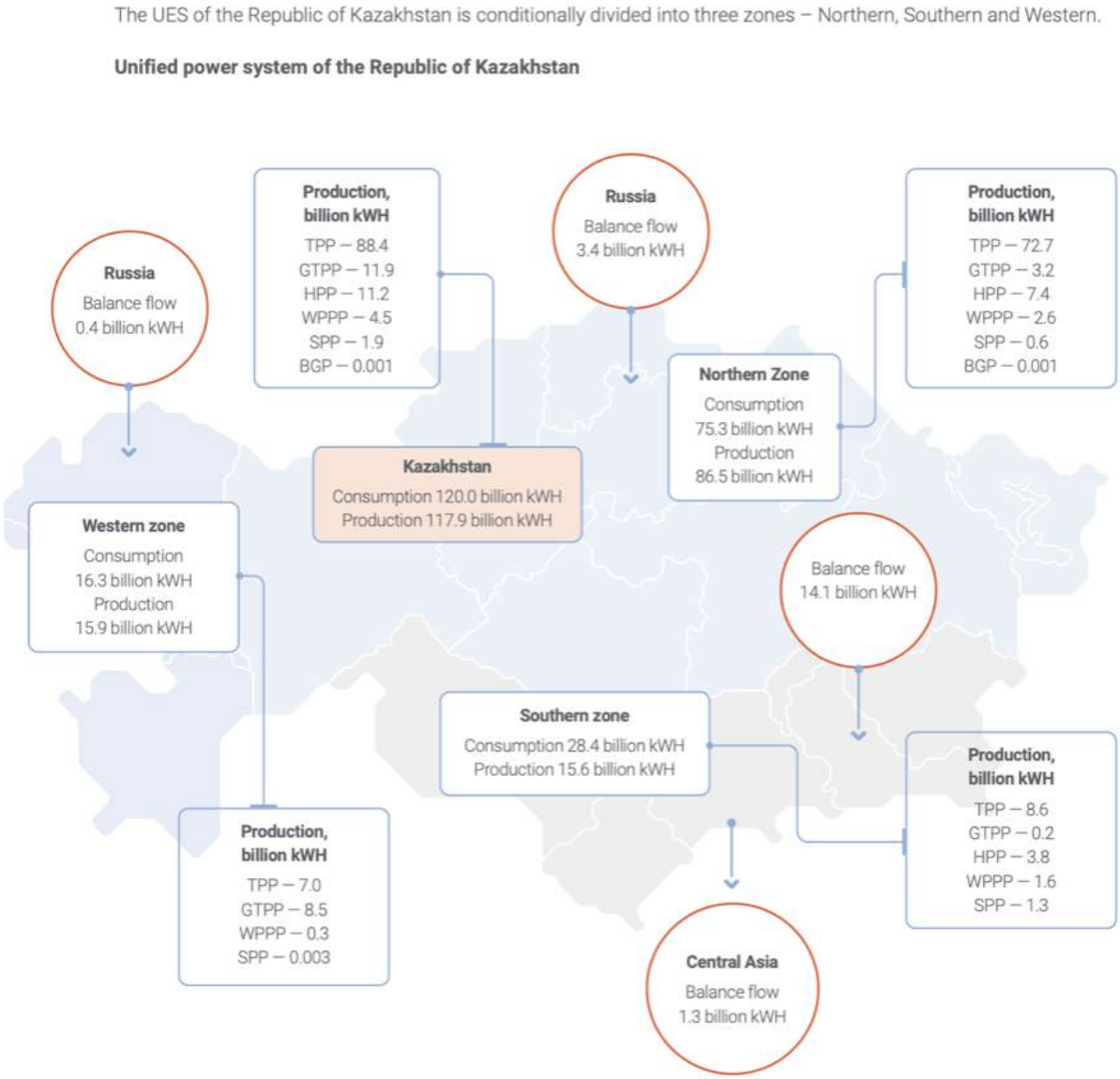


Figure 4. *Infographics of Kazakhstan's Energy Balance in 2024.* This infographic shows the structure of Kazakhstan's energy balance in 2024, illustrating the relative contribution of different energy sources to national electricity generation. Source: Samruk-Energy JSC, *Infographics of Kazakhstan's Energy Balance in 2024*.

3. Methodology

3.1 Research Design

This research implements multi-criteria evaluation for identifying the appropriate nuclear reactor technology. This approach is selected as it facilitates the analysis of each technology from multiple perspectives, thereby providing a stronger basis for an objective conclusion of the research. Evaluation built around one criterion tends to be limited and insufficient; individual analysis of technology's electric output will omit the safety and security concerns and vice versa. Moreover, the same methodology was employed in previous studies on a similar topic. Kim devised criteria for the future nuclear power plant;

Pospelov et al. have implemented technical criteria to ensure the fairness of the evaluation; Matebay considered environmental, economical, technical, financial factors to determine the most suitable nuclear reactor type. This establishes multi-criteria evaluation as a reliable method to conduct a present study.

3.2 Literature Review and Criteria Development

Literature reviews present an accessible and broad way of gathering information from existing papers, reports and documentation. It was used in the present research to provide the basis for multi-criteria evaluation and for the context of Kazakhstan specific environmental and national circumstances. In accordance with the literature review, the following six criteria were developed: safety and security, climate and water resources, grid compatibility, economic costs, fuel cycle and waste management, institutional fit.

3.3 Criteria Weightings

It is recognized that each factor carries a different degree of importance regarding choosing a suitable nuclear reactor technology. To account for this, individual weightings were attributed to each of the criteria. Criteria and their respective weightings are shown in Table 1.

Table 1. *Evaluation criteria and weighting used to assess selected reactor designs for Kazakhstan's first nuclear power plant.* This table shows the weighting of each criterion in the final reactor suitability assessment. The total weighting equals 100%.

Criterion	Weight
Safety and Security	25%
Climate and Water Resources	20%
Grid Compatibility	15%
Economical costs	15%
Fuel Cycle and Waste Management	12.5%
Institutional Fit	12.5%

Safety and security, which were assigned 25% weight, are prioritized due to the detrimental implications associated with a former nuclear testing site near Semipalatinsk, Eastern Kazakhstan. Studies conducted by Kuralbayeva et al. and Mussin et al. demonstrate how radioactive contamination has led to negative health-related and environmental consequences. Kazakhstan's previous experience with nuclear energy has become a sensitive topic, which to date is widely discussed in public, highlighting the importance of safety and security concerns.

Climate and water resources criteria received 20% weight due to the future site of Kazakhstan's nuclear power plant: near Ulken, coast of Lake Balkhash. Nuclear power plants utilize water resources heavily, which raises potential water unavailability concerns. Concerns become even more relevant when put in the wider context of Kazakhstan's water sustainability: Alimkulov et al. and Yang et al. predict that the water scarcity issue will arise by 2040. Hence, water usage is a significant factor to consider in selection of nuclear reactor technology.

Criteria such as economic cost and grid compatibility were each assigned 15% weight. Both assess the practicality of implementing a nuclear reactor technology. Grid compatibility determines whether a nuclear power plant is able to be successfully integrated into Kazakhstan's energy system. This includes the power generation capacity, transmission requirements and single-contingency risks. Economic cost determines the feasibility of the nuclear power plant project, assessing the potential financial risks and possible delays. Both of the criteria are weighted less than safety and water resources concerns because the risks are partially easier to mitigate. However, grid compatibility and economic cost remain significant factors (Kim).

Fuel cycle with waste management and institutional fit were assigned 12.5% weight. Discussed by Matebay, fuel cycle and waste management were included due to Kazakhstan's existing uranium production capacity, while institutional fit is important due to the fact that this nuclear power plant is going to be the first in the country. The difference between 15% and 12.5% reflects that these criteria are significant but they largely depend on supplier agreements and long-term policies, which are relatively more changeable compared to other criteria and factors.

3.4 Scoring Rubrics

Each criterion is assigned a score from 1 to 5; 1 indicates a poor suitability, 5 represents a great fit for Kazakhstan specific requirements. In order to systemize and organize the scoring, main sub-indicators for each of the respective criteria are developed. They will provide a basis for the scoring, ensuring objective evaluation. Criteria with sub-indicators are presented in Table 2.

Table 2. *Evaluation Criteria and Main Sub-Indicators Used to Assess Selected Reactor Designs for Kazakhstan's First Nuclear Power Plant.* This table outlines the key factors used to assess each reactor design across the six evaluation criteria.

Criterion	Main sub-indicators
Safety and security	Operating record, containment design, passive/active safety systems, seismic resistance
Climate and water resources	Cooling-water demand, thermal discharge risk, tolerance to temperature extremes
Grid compatibility	Unit size, reserve requirement, load-following ability, transmission constraints
Economic costs	Overnight construction cost, construction time, cost-overrun risk, financing terms, tariff pressure
Fuel cycle and waste management	Fuel supply security, enrichment dependence, fuel fabrication compatibility, spent fuel strategy, waste management
Institutional fit	Vendor export record, technology transfer, geopolitical and public acceptance risks

3.5 Final Score

A final score was calculated using the scoring and weighting system outlined previously. For each technology, the score was multiplied by the respective criterion's weighting. This procedure is repeated for all of the scores of an individual technology. Summing up the results of such operations will produce a final score for the technology, ranging from 1 to 5. These final scores will be central to the present study's discussion, providing a pivotal point for the selection of the appropriate nuclear reactor technology.

4. Results

4.1 Safety and Security

Safety and security criteria were assessed based on both operational experience and technical characteristics of nuclear reactors due to the highest weighting. The evaluation is presented in Table 3.

Table 3. *Safety and Security Assessment of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant.* This table compares the deployment status, risk indicators, seismic resistance, and key safety features of each reactor design used to determine its safety and security score.

Reactor	Deployment status	Risk indicators	Seismic resistance	Key safety features	Score
VVER-1200	● 6 operating ● 13 under construction ● 6 planned	CDF ~10^{-7}	~0.25g	● Double containment ● Active/passive safety ● Core melt localization ● Passive heat removal	4/5
HPR1000	● 6 operating ● 14 under construction	CDF <10^{-6}; LRF <10^{-7}	0.3g	● Double-shell containment ● Active/passive safety ● Passive residual heat removal ● Aircraft-crash resistance	4/5
EPR1200	● 0 operating ● 0 under construction	CDF ~10^{-7}; core-meltdown frequency <10^{-5}	0.3g	● Three safety trains ● Double containment ● Core catcher ● Catalytic hydrogen recombiners	4/5
BWRX-300	● 0 operating ● first unit under construction	CDF <10^{-7}; LRF <10^{-8}	0.3g	● Natural circulation ● Three ICS trains ● 72-hour DC power ● No emergency AC diesels required for reactor safety	4/5

Note. CDF = core damage frequency; LRF = large release frequency; ICS = isolation condenser system; AC = alternating current; DC = direct current; g = acceleration due to gravity. Scores are based on the 5-point assessment scale described in the methodology section.

Safety results are closely clustered, with all four reactors receiving 4/5, indicating that safety performance is not the main differentiating criterion. The main indicators are CDF, LRF, and seismic resistance. CDF estimates the probability of severe reactor-core damage per reactor-year; for example, 10^{-7} means a modelled probability of about one event per ten million reactor-years. LRF estimates the probability of a major radioactive release after severe accident progression, so lower values indicate stronger containment and severe-accident protection. HPR1000, EPR1200, and BWRX-300 report 0.3 g seismic resistance, while VVER-1200 is reported at approximately 0.25 g (Matebay; Xing et al.). BWRX-300 has the lowest reported LRF, below 10^{-8}/reactor-year, but its lack of operating experience limits its score (GE Hitachi Nuclear Energy; GE Vernova Hitachi Nuclear Energy Americas). EPR1200 has strong severe-accident features, including a core catcher and catalytic hydrogen recombiners, but the specific EPR1200 variant remains unbuilt (EDF). Therefore, the main distinction is the balance between safety design maturity and deployment evidence.

4.2 Climate and Water Resources

Due to the potential water scarcity in Kazakhstan, optimized usage of water resources by reactors is prioritized. Therefore, performance of each nuclear reactor technology under cooling-related conditions is analyzed, as illustrated in Table 4.

Table 4. *Water-Use and Cooling Suitability Assessment of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant*. This table summarizes the thermal power, efficiency, cooling options, and key water-risk indicators used to assign each reactor design its score under the climate and water resources criterion.

Reactor	Thermal power	Efficiency	Cooling options / water data	Key water-risk indicator	Score
VVER-1200	~3212 MWt	~37%	Closed-cycle cooling expected at Ulken/Balkhash	● Two units estimated to cause ~60 million m³/year non-return evaporative losses	3/5
HPR1000	~3050 MWt	~37%	Once-through or closed-cycle cooling possible	● UK reference case: 198,000 m³/h cooling water; ~9.7°C temperature rise	3/5
EPR1200	~3300 MWt	37%	Open or closed cooling; advanced cooling-tower design	● Large water-cooled PWR ● Substantial evaporative losses under closed-cycle cooling	3/5
BWRX-300	870 MWt	Lower total heat burden	Dry, hybrid, mechanical-draft, or once-through cooling	● Maximum recommended condenser inlet temperature: 37.8°C	4/5

Note. MWt = megawatts thermal; m³/year = cubic metres per year; m³/h = cubic metres per hour. Scores are based on the 5-point assessment scale described in the methodology section.

Climate and water performance is mainly determined by reactor size, thermal power, and cooling method. MWt measures heat produced in the reactor, while MWe measures electricity delivered to the grid. Since large PWRs convert only about 37% of thermal energy into electricity, the remaining heat must be rejected through cooling systems. This explains why VVER-1200, HPR1000, and EPR1200 remain water-sensitive: all have thermal powers above 3000 MWt. HPR1000 has the clearest cooling-water figure, with 198,000 m³/h in the UK once-through reference case and an estimated 9.7 °C temperature rise (Matebay). VVER-1200 has the most Kazakhstan-specific water-risk figure, with two units estimated to cause around 60 million m³/year of non-return evaporative losses (Alimkulov et al.; Nosko). BWRX-300 performs better because its thermal power is only 870 MWt, reducing total heat rejection, and because dry or hybrid cooling options could reduce freshwater withdrawal (GE Vernova Hitachi Nuclear Energy Americas; General Nuclear System Ltd.). However, dry cooling may increase capital cost and reduce efficiency, so even the SMR option requires site-specific environmental assessment.

4.3 Grid Compatibility

Grid compatibility of the reactor designs was evaluated based on how flexible power generation of each reactor is and what stress on Kazakhstan's energy system it generates, as demonstrated in Table 5.

Table 5. *Grid Compatibility Assessment of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant.* This table compares the net electrical output, single-unit contingency, and load-following capability used to assign each reactor design its score under the grid compatibility criterion.

Reactor	Net electrical output	Single-unit contingency	Load-following capability	Score
VVER-1200	~1170–1200 MWe	>1 GW	Available	4/5
HPR1000	~1080–1180 MWe	>1 GW	Available	4/5
EPR1200	1200 MWe	~1.2 GW	25–100% in 30 min	4/5
BWRX-300	~300 MWe	~300 MW	50–100% at 0.5%/min	5/5

Note. MWe = megawatts electric; MW = megawatts; GW = gigawatts. Single-unit contingency refers to the amount of generating capacity lost if one reactor unit unexpectedly shuts down. Scores are based on the 5-point assessment scale described in the methodology section.

Grid compatibility depends mainly on unit size and operational flexibility. Single-unit contingency refers to the amount of electricity suddenly lost if one reactor shuts down unexpectedly. For Kazakhstan, this is important because losing a 1.1–1.2 GWe unit would place much greater pressure on reserve capacity and transmission infrastructure than losing a 300 MWe unit. This explains why BWRX-300 receives the highest score: its smaller size reduces contingency risk and allows gradual capacity additions. Among the large reactors, HPR1000 has a slight advantage because its 1080–1180 MWe output is somewhat lower than VVER-1200 and EPR1200 (Matebay). Load-following data also matter because they show how well reactors can adjust output when electricity demand or renewable generation changes. EPR1200 can ramp from 25% to 100% power in 30 minutes, while BWRX-300 can load-follow from 50% to 100% at 0.5% per minute (EDF; GE Vernova Hitachi Nuclear Energy Americas). Therefore, all large reactors can support baseload generation, but BWRX-300 is structurally better suited to a fragmented and developing grid.

4.4 Economic and Financial Costs

Construction of the nuclear power plant generates significant pressure for the country's economy. Hence, each of the reactor designs is analyzed in the context of optimal financing terms and on time deployment, as shown in Table 6.

Table 6. *Economic and Financial Cost Assessment of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant.* This table summarizes the cost evidence, construction-duration evidence, and financing evidence used to assign each reactor design its score under the economic costs criterion.

Reactor	Cost evidence	Construction-duration evidence	Financing evidence	Score
VVER-1200	● Kazakhstan estimate: USD 14–15 billion for two units; ● Rooppur: USD 12.65 billion for two units	● 7.5 - 10 years	● Rooppur: 90% Russian state loan, 4% interest, 10-year grace period	4/5
HPR1000	No Kazakhstan-specific cost estimate identified	● Fuqing Unit 5: 5.7 years ● Karachi units: 5.5–5.7 years	● Pakistan loans: 1%, 2%, and 6% interest rates ● Atucha 3 proposal: USD 6.7 billion at 4.5%	4/5
EPR1200	EPR reference projects: ● Olkiluoto 3: €3.2bn→€11bn; ● Flamanville: €3.3bn→€23.7bn; ● Hinkley Point C: €20.8bn→~€55bn	● Olkiluoto 3: 16.5 years; ● Flamanville 3: 17 years	● Olkiluoto 3: €570m Coface loan guarantee	2/5
BWRX-300	● Darlington: USD 15.4 billion for 1200 MW; first unit around USD 19.1 million/MW	● Below 30 months	● USD 713 million Canada Infrastructure Bank loan. ● USD 2.1–2.2 billion public equity support	3/5

Note. USD = United States dollars; € = euros; MW = megawatts; bn = billion; m = million. Scores are based on the 5-point assessment scale described in the methodology section.

Economic performance is assessed through four comparable indicators: capital-cost evidence, construction duration, financing terms, and the reliability of the cost evidence. HPR1000 scored 4/5 as HPR1000 has stronger recent construction-duration data, including 68.7 months for Fuqing Unit 5 and 5.5–5.7 years for Karachi (Kwak; "Federal Loan Provides Certainty"). VVER-1200 also received 4/5 as it has more direct Kazakhstan-related cost evidence and a relevant export financing model through Rooppur (Akter and Islam; Matebay). EPR1200 receives the lowest score because it has no direct construction record and must be assessed through larger EPR projects, which show major cost escalation and delays (Gaudiaut). BWRX-300 sits between these cases. Darlington provides real first-of-a-kind cost evidence, but current costs are not yet representative of mature nth-of-a-kind SMR economics ("Canada's First SMR Project"; GE Vernova Hitachi Nuclear Energy Americas).

4.5 Fuel Cycle and Waste Management

Given the existing uranium production capacity in Kazakhstan, technical characteristics of fuel used by each nuclear reactor design is analyzed. Also, former historical experience with nuclear energy (Semipalatinsk) reflects the importance of reprocessing waste. These criteria are assessed in Table 7.

Table 7. *Fuel Cycle and Waste Management Assessment of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant.* This table summarizes the fuel type, enrichment requirements, fuel-cycle length, burnup, supply-chain capacity, and waste-management status used to assign each reactor design its score under the fuel cycle and waste management criterion.

Reactor	Fuel type	Enrichment	Fuel cycle	Burnup	Conversion capacity	Enrichment capacity	Fuel-fabrication capacity	Reprocessing / waste status	Score
VVER-1200	UO_2	Up to 4.95%	12–18 months	Up to 70 MWd/kgU	12,500 tU/year	27.7 million SWU/year	2,760 t/year	700-800 t/year	3/5
HPR1000	UO_2	Around 5%	18 months	55–60 GWd/MTU	15,000 tU/year	7.5–11 million SWU/year	1,400 t/year	Developing	4/5
EPR1200	UO_2	Standard LEU range	Up to 24 months	Not specified	15,000 tU/year	7.5 million SWU/year	3,250 t/year	1,700 t/year	4/5
BWRX-300	UO_2	3.81% average / 4.95% maximum	12–24 months	About 50 GWd/MTU	Not specified	No HALEU required	Not specified	Not specified	3/5

Note. UO_2 = uranium dioxide; LEU = low-enriched uranium; HALEU = high-assay low-enriched uranium; MWd/kgU = megawatt-days per kilogram of uranium; GWd/MTU = gigawatt-days per metric ton of uranium; SWU/year = separative work units per year; tU/year = tonnes of uranium per year; t/year = tonnes per year. Scores are based on the 5-point assessment scale described in the methodology section.

Fuel-cycle performance is assessed through standardized front-end and back-end indicators: fuel type, enrichment, fuel-cycle length, burnup, conversion capacity, enrichment capacity, fuel-fabrication capacity, and reprocessing or waste status. Enrichment refers to the percentage of uranium-235 in nuclear fuel, while burnup measures the amount of energy extracted from fuel before discharge. Higher burnup can improve fuel efficiency, but it may also make spent-fuel management more complex. The table shows that VVER-1200, HPR1000, and EPR1200 are supported by more complete supplier-side fuel-cycle systems than BWRX-300. Russia has the largest enrichment capacity at 27.7 million SWU/year, while China and France show strong conversion capacity at around 15,000 tU/year each (Matebay). France has the strongest identified reprocessing capacity through La Hague at 1,700 t/year, compared with Russia's RT-1 capacity of 400 t/yearand planned RT-2 expansion (EDF; Matebay). BWRX-300 uses conventional low-enriched UO_2 fuel and avoids HALEU, but comparable supplier-level conversion, enrichment, and fabrication data are not specified in the available sources. Therefore, the main distinction is between integrated supplier fuel-cycle support and Kazakhstan's unresolved long-term spent-fuel strategy, which remains a limitation for all technologies (GE Vernova Hitachi Nuclear Energy Americas).

4.6 Institutional Fit

Kazakhstan has no nuclear power plants operating in the country currently. This highlights the significance of institutional fit when selecting an appropriate nuclear reactor design. This criteria is evaluated based on the previous operational record of the vendor companies, as shown in Table 8.

Table 8. *Institutional Fit Assessment of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant.* This table summarizes the operating experience, construction pipeline, and main institutional limitations used to assign each reactor design its score under the institutional fit criterion.

Reactor	Operating units	Construction pipeline	Main limitation	Score
VVER-1200	6	● 13 under construction ● 6 planned	● Geopolitical dependence ● Public opposition to Rosatom	4/5
HPR1000	6	● 14 under construction	● Lower public preference for China: 16%	4/5
EPR1200	0	● 0 under construction	● Unbuilt design ● EPR reference projects delayed and over budget	3/5
BWRX-300	0	● First unit under construction	● First-of-a-kind licensing, supply-chain, and operator-training risks	3/5

Note. Operating units and construction pipeline indicate the current level of international deployment experience for each reactor design. Scores are based on the 5-point assessment scale described in the methodology section.

Delivery and institutional fit are assessed through operating units, construction pipeline, Kazakhstan or export relevance, and institutional limitations. Operating units are important because they reduce uncertainty in licensing, construction, supply-chain management, operator training, and long-term maintenance. VVER-1200 and HPR1000 score highest because both have 6 operating units and active construction pipelines. VVER-1200 has the strongest Kazakhstan-specific momentum because Rosatom has been selected as consortium leader and Ulken site work has begun, but this is balanced by geopolitical dependence and public opposition to Rosatom (Tengrinews.kz, "Who Will Build"). HPR1000 has strong export credibility through Pakistan and broader Kazakhstan–China economic links, although public preference for China is lower (Matebay). EPR1200 and BWRX-300 score lower because neither has an operating unit in the exact evaluated design. EPR1200 remains unbuilt, while BWRX-300 is still a first-of-a-kind project expected around 2030 (GE Vernova Hitachi Nuclear Energy Americas; "Federal Loan Provides Certainty").

4.7 Final Weighted Scores

In accordance with the methodology of the current study, the final weighted scores for each nuclear reactor design are calculated using criteria and the respective weighting. These scores are presented in Table 9.

Table 9. *Final Weighted Suitability Scores and Ranking of Selected Reactor Designs Considered for Kazakhstan's First Nuclear Power Plant.* This table presents the criterion-level scores and final weighted total scores used to rank the reactor designs in the overall suitability assessment.

Rank	Reactor	Safety	Climate / water	Grid	Economic	Fuel / waste	Institutional	Total
1	HPR1000	4	3	4	4	4	4	3.80 / 5.00
2	BWRX-300	4	4	5	3	3	3	3.75 / 5.00
3	VVER-1200	4	3	4	4	3	4	3.68 / 5.00
4	EPR1200	4	3	4	2	4	3	3.38 / 5.00

Note. Criterion scores are based on the 5-point assessment scale described in the methodology section. Total scores are weighted according to the evaluation criteria weights shown in Table 1.

5. Discussion

5.1 Discussion of Results

The results indicate that no reactor technology is universally superior across all criteria. HPR1000 achieved the highest weighted score, 3.80/5.00, followed closely by BWRX-300 at 3.75/5.00, VVER-1200 at 3.68/5.00, and EPR1200 at 3.38/5.00. The first three technologies share a gap of only 0.12 points, which suggests that the ranking should be interpreted as a conditional assessment rather than a decisive prediction. The final selection of the appropriate nuclear reactor technology depends on Kazakhstan's policy priorities, such as ecological sustainability, efficiency maximization, and financing security.

All of the reactor technologies selected for the evaluation scored 4/5 in the safety and security criterion. This shows that modern nuclear reactor technologies have advanced safety measures, including seismic resistance, severe-accident mitigation features, and low core damage frequencies. Hence, reactor technologies were not significantly differentiated by this criterion. None of the reactors received the highest score due to minor individual disadvantages. VVER-1200 was reported to have lower seismic resistance; HPR1000 relies only on Chinese-side evidence; and EPR1200 and BWRX-300 do not have any operational unit record (EDF; GE Hitachi Nuclear Energy; GE Vernova Hitachi Nuclear Energy Americas; Matebay; Xing et al.).

HPR1000 received a score of 3.80, ranking first among all reactor technologies selected for the multi-criteria evaluation. This demonstrates that HPR1000 is the most balanced option. It received 4/5 in safety and security, grid compatibility, economic costs, fuel cycle, and institutional fit. Its moderate unit size, comfortable financing conditions, and reliable operational record in Pakistan cement HPR1000 as a suitable design for Kazakhstan's first nuclear power plant (Matebay; Xing et al.). Despite this, the use of a large water-cooled PWR with substantial heat-rejection requirements is a significant limitation of the HPR1000 reactor design. Accordingly, it received a score of 3/5 in the climate and water resources criterion.

BWRX-300 comes in second with a score of 3.75. It is the strongest Kazakhstan-specific reactor selection because it directly addresses water stress and grid fragmentation concerns. Due to its compact design and smaller power generation of 300

MWe, the single-contingency risk is reduced. Its 870 MWt thermal power lowers cooling demand, decreasing water stress on Lake Balkhash. These advantages justify the 5/5 score received by BWRX-300 in grid compatibility and 4/5 in climate and water resources. However, it received 3/5 in economic costs, fuel cycle, and institutional fit because it has no operational record, with the first project currently under construction. Due to this, BWRX-300 lacks reliable fuel-cycle and deployment record data (GE Vernova Hitachi Nuclear Energy Americas; "Federal Loan Provides Certainty"). Therefore, BWRX-300 may be selected when long-term sustainability is prioritized, but it poses significant short-term economic costs and institutional burdens.

VVER-1200 is ranked third with a score of 3.68. It remains a practical but geopolitically conditional option. VVER-1200's main advantages include a large operational record, technology localization, and potentially favorable financing comparable to the Rooppur model (Akter and Islam; Tengrinews.kz, "Who Will Build"). However, its disadvantages include significant water demand and geopolitical risk, which may lead to potential deployment delays. Additionally, the 4/5 economic score is highly conditional. If Kazakhstan receives concessional financing, VVER-1200 remains competitive; under less favorable tariff or debt structures, its score would likely decrease to 3/5 (Markopoulos and Karkazis).

EPR1200 ranks last with a score of 3.38. This is mainly due to economic and delivery risk. Technically, it has strong safety features and benefits from France's advanced fuel-cycle infrastructure. However, EPR1200 itself has no operating record, while reference EPR projects have experienced major cost overruns and delays, including Olkiluoto 3, Flamanville 3, and Hinkley Point C (Gaudiaut; Matebay). Additionally, compared to other alternatives, EPR1200 has the greatest construction cost. For Kazakhstan, a country with no nuclear facilities in operation, this creates implementation risk, explaining its 2/5 economic score and 3/5 institutional score.

In general, HPR1000 is the most suitable option under the current framework because it performs consistently across most criteria. However, the result is dependent on Kazakhstan's priorities. If immediate implementation is prioritized, VVER-1200 remains the best alternative. If water conservation and grid flexibility are prioritized, BWRX-300 becomes more attractive. If regulatory credibility and technical reliability are prioritized, EPR1200 has strategic value but appears less suitable as the primary reactor choice.

5.2 Limitations and Next Steps

This study has several limitations. Before the discussion of limitations, it is important to clarify that the aim of the study is not to provide a fully detailed technical reactor evaluation, the study's purpose was to provide a comparative overview of selected reactor designs for Kazakhstan's first nuclear power plant. Therefore, current study is intended to support preliminary evaluation rather than replacing engineering reviews or financial feasibility studies. The main impact of the current study is organizing technical, environmental, institutional, economic and ecological implications of the selected reactor designs into a structured overview, highlighting the advantages and disadvantages of each reactor design.

To start with, data availability was uneven. Established and broadly recognized reactor technologies such as VVER-1200 and HPR1000 had more complete information on construction and technical details, which facilitated the in-depth analysis. In contrast, more recent reactor technologies such as BWRX-300 and EPR1200 had limited data because neither design has any operating units as of now. Reference projects, such as Darlington for BWRX-300, and similar reactors, such as EPR1650, were used in order to compensate for unavailable information. Even though this method provides a reasonable approximation, it introduces uncertainty to the multi-criteria evaluation.

The second limitation of the current study is the weighting system. It provides a simplification of a complex national decision. Although it is true that factors contribute to the selection of reactor technology differently, the assigned weight of each individual factor remains subjective. Stakeholders may prioritize different factors, which directly influences the importance and weighting of the respective factors. For example, a higher water-resource weighting would favor the BWRX-300 reactor design, while a higher economic cost weighting would prioritize HPR1000 or VVER-1200 reactors.

The third limitation of the current study is the scoring system. The 1–5 scoring scale oversimplifies the complex assessment of reactor technologies. This scale allows for effective and clear differentiation between reactor designs. However, it is challenging to reflect smaller discrepancies using a 1–5 scoring scale. This leads to different reactors receiving the same score, which makes them seem equal in a certain category, while in fact they are not.

Future research should implement different methodologies in order to ensure a more objective and detailed analysis. Specifically, sensitivity analysis, including the Analytic Hierarchy Process (AHP) or Monte Carlo simulation, can be used to test how different weights affect the final ranking and to account for stakeholder preferences. Additionally, it is advised to include updated information on the construction of Kazakhstan's first nuclear power plant. This will allow for a more relevant and accurate evaluation.

6. Conclusion

This study implemented a multi-criteria evaluation of VVER-1200, HPR1000, EPR1200, and BWRX-300 for Kazakhstan's first nuclear power plant. The study demonstrates that HPR1000 is the most appropriate option under the selected criteria, achieving the highest weighted score of 3.80/5.00, followed closely by BWRX-300 at 3.75/5.00 and VVER-1200 at 3.68/5.00. EPR1200 ranked lowest at 3.38/5.00 due to economic and delivery risks. Overall, the findings suggest that Kazakhstan's reactor choice should not be based only on technical safety, since all designs selected for evaluation performed well in this respective area. Instead, the decision should also include such factors as water availability, grid compatibility, financing terms, fuel-cycle dependence, and institutional feasibility. HPR1000 appears to offer the most balanced option, while BWRX-300 may become a better alternative if Kazakhstan prioritizes water conservation and grid adaptability. Future research should use sensitivity analysis, site-specific water modelling, and updated construction information to strengthen the reactor-selection assessment.

Works Cited

Akter, S., and M. S. Islam. "Cost Modeling and Policy Insights for Deploying Two VVER-1200 Reactors in the Newcomer Nuclear Country of Bangladesh." *Energy Strategy Reviews*, vol. 64, 2026, article 102102. https://doi.org/10.1016/j.esr.2026.102102.

Alimkulov, S., et al. "Long-Term Water Level Projections for Lake Balkhash Using Scenario-Based Water Balance Modeling under Climate and Socioeconomic Uncertainties." *Water*, vol. 17, 2025, article 2021. https://doi.org/10.3390/w17132021.

Asmolov, V. G., et al. "New Generation First-of-the-Kind Unit: VVER-1200 Design Features." *Nuclear Energy and Technology*, vol. 3, no. 4, 2017, pp. 260–69. https://doi.org/10.1016/j.nucet.2017.10.003.

Batyrbekov, E., et al. "Conceptual View of the Implementation of the Nuclear Energy Program in the Republic of Kazakhstan." *Energies*, vol. 17, no. 22, 2024, article 5788. https://doi.org/10.3390/en17225788.

Bespalyy, S., and Y. Bespalaya. "Balancing Electricity and Integrating Renewable Energy Facilities into the Unified Energy System of Kazakhstan: Challenges and Opportunities." *E3S Web of Conferences*, vol. 623, 2025, article 03001. https://doi.org/10.1051/e3sconf/202562303001.

Burck, Jan, et al. *The Climate Change Performance Index 2026: Results*. Germanwatch, NewClimate Institute, and Climate Action Network International, 18 Nov. 2025. https://www.germanwatch.org/en/93330.

"BWRX-300 Small Modular Reactor." *GE Vernova Hitachi Nuclear*, accessed 18 May 2026. https://www.gevernova.com/nuclear/carbon-free-power/bwrx-300-small-modular-reactor.

"Canada's First SMR Project: How Is CAD20.9 Billion Cost Calculated?" *World Nuclear News*, 23 May 2025. https://world-nuclear-news.org/articles/what-is-the-budget-for-canadas-first-smr-project.

EDF. "Our Reactor Portfolio." *EDF Poland*. https://poland.edf.com/en/our-activities/nuclear/edf-a-unique-nuclear-player/our-reactor-portfolio. Accessed 24 June 2026.

"Electricity and Energy Storage." *World Nuclear Association*, accessed 12 May 2026. https://world-nuclear.org/information-library/current-and-future-generation/electricity-and-energy-storage.

ENERGY Insight & Analytics. "Resilient Energy System: Energy Storage, Smart Grids, and Efficiency." *Analytical Center "ENERGY" LLP*, 1 Dec. 2025. https://s3-prod.exia.kz/articles/251201_Article_28_Resilent_Energy_System_EN.pdf.

"Energy Storage Systems Will Be Built in Kazakhstan." *Kapital.kz*, 25 Feb. 2026. https://kapital.kz/economic/145352/sistemy-nakopleniya-energii-budut-stroit-v-kazahstane.html.

"Federal Loan Provides Certainty for Canadian SMR Build." *World Nuclear News*, 25 Oct. 2022. https://world-nuclear-news.org/articles/federal-loan-provides-certainty-for-canadian-smr-b.

"Fuel and Energy Balance of the Republic of Kazakhstan (2024)." *Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan*, accessed 12 May 2026. https://stat.gov.kz/en/industries/business-statistics/stat-energy/publications/411630/.

Fyza, N., A. Hossain, and R. Sarkar. "Analysis of the Thermal-Hydraulic Parameters of VVER-1200 Due to Loss of Coolant Accident Concurrent with Loss of Offsite Power." *Energy Procedia*, vol. 160, 2019, pp. 155–61. https://doi.org/10.1016/j.egypro.2019.02.131.

Gaudiaut, T. "European Nuclear Projects Face Major Cost Overruns and Delays." *Statista*, 24 Apr. 2026. https://www.statista.com/chart/36138/cost-overruns-and-construction-delays-of-epr-nuclear-reactors-in-europe/.

GE Hitachi Nuclear Energy. *Licensing Topical Report: BWRX-300 Reactor Pressure Vessel Isolation and Overpressure Protection*. NEDO-33910, rev. 0, GE-Hitachi Nuclear Energy Americas, Dec. 2019. https://www.nrc.gov/docs/ML1936/ML19364A212.pdf.

GE Vernova Hitachi Nuclear Energy Americas LLC. *BWRX-300 General Description*. 005N9751, rev. H, Oct. 2025. https://www.gevernova.com/content/dam/gevernova-nuclear/global/en_us/documents/carbon-free-power/005N9751-BWRX-300-General-Description.pdf.

General Nuclear System Ltd. *UK HPR1000 GDA Project: Pre-Construction Environmental Report, Chapter 8: Conventional Impact Assessment*. HPR/GDA/PCER/0008, rev. 000, 2018. https://ukhpr1000.co.uk/wp-content/uploads/2018/11/HPR-GDA-PCER-0008-Pre-Construction-Environmental-Report-Chapter-8-Conventional-Impact-Assessment.pdf.

Gonyeau, Joseph. "PWR Cycle." *The Virtual Nuclear Tourist*, 2 Nov. 2002. https://www.nucleartourist.com/type/pwr_cycle.htm.

Hruskovic, J. *Comparative Review on Safety Systems of VVER Reactors*. https://www.kns.org/files/pre_paper/54/25A-075-Jan.pdf.

International Atomic Energy Agency. *Small Modular Reactors: Advances in SMR Developments 2024*. IAEA, 2024. https://www.iaea.org/publications/15790/small-modular-reactors-advances-in-smr-developments-2024.

Insepov, Zinetula, et al. "Small Modular Nuclear Power Reactors as a Driver of Development of Nuclear Technologies." *Energies*, vol. 18, no. 21, 2025, article 5766. https://doi.org/10.3390/en18215766.

KEGOC. "National Power System." Accessed 11 May 2026. https://www.kegoc.kz/en/electric-power/natsionalnaya-energosistema/.

Kim, Alexandr. *Comparative Analysis on Decision Making in the Case of Nuclear Power Plant Development in the Republic of Kazakhstan*. Seoul National University, 2019. *DBpia*, https://www.dbpia.co.kr/journal/detail?nodeId=T15113420.

Kuralbayeva, K., et al. *Long-Term Health Effects of Nuclear Tests: The Semipalatinsk Case*. 2025. https://www.econstor.eu/handle/10419/309303.

Kwak, S. *An Analysis of Key Economic Drivers in Nuclear Energy and Implications for Optimal Plant Design*. Master's thesis, Seoul National University, 2025. https://s-space.snu.ac.kr/handle/10371/228388.

Markopoulos, T., and J. Karkazis. *Akkuyu Nuclear Power Plant: Is It Possible? An Updated and Critical Analysis*. Center of Strategic Studies, Aug. 2025. https://doi.org/10.13140/RG.2.2.36174.98882/1.

Matebay, B. *Vendor and Technology Selection Recommendations for Kazakhstan's First Nuclear Power Plant*. Master's thesis, The University of Texas at Austin, 2025. https://repositories.lib.utexas.edu/items/0f4f3fb8-aee0-4b54-b068-eb996b7c69a6.

"The Ministry of Energy Plans to Implement 68 Projects to Introduce New 6.7 GW of Capacity by 2029." *Official Information Source of the Prime Minister of the Republic of Kazakhstan*, accessed 11 May 2026. https://primeminister.kz/en/news/the-ministry-of-energy-plans-to-implement-68-projects-to-introduce-new-67-gw-of-capacity-by-2029-30658.

Mussin, B., et al. "Sociological Analysis of Public Opinion on the Construction of a Nuclear Power Plant in Kazakhstan." *Polityka Energetyczna – Energy Policy Journal*, vol. 28, no. 1, 2025, pp. 151–72. https://doi.org/10.33223/epj/199840.

"Nuclear Fission Chain Reaction." *Nuclear Power*. https://www.nuclear-power.com/nuclear-power/reactor-physics/nuclear-fission-chain-reaction/. Accessed 24 June 2026.

"Nuclear Power Reactors." *World Nuclear Association*, 23 Mar. 2026. https://world-nuclear.org/information-library/nuclear-power-reactors/overview/nuclear-power-reactors.

Nosko, Artyom. "How Much Water Will Kazakhstan's First Nuclear Power Plant Need for Cooling?" *Orda.kz*, 14 May 2026. https://en.orda.kz/how-much-water-will-kazakhstans-first-nuclear-power-plant-need-for-cooling-10405/.

"Over 1 Trillion Tenge to Be Allocated for Modernization of Energy and Utilities Infrastructure." *Official Information Source of the Prime Minister of the Republic of Kazakhstan*, accessed 12 May 2026. https://primeminister.kz/en/news/over-1-trillion-tenge-to-be-allocated-for-modernization-of-energy-and-utilities-infrastructure-31348.

"Phased Introduction of New Capacities to Ensure the Country's Energy Security." *gov.egov.kz*, accessed 12 May 2026. https://betaegov.kz/memleket/entities/energo/press/news/details/1015553.

"Physics of Uranium and Nuclear Energy." *World Nuclear Association*, accessed 17 Apr. 2026. https://world-nuclear.org/information-library/nuclear-fuel-cycle/introduction/physics-of-nuclear-energy.

Pospelov, V. A., V. V. Baklanov, and P. G. Korovikov. "Comparative Analysis of Reactor Technologies." *Bulletin of the NNC RK*, vol. 1, no. 4, 2019, pp. 73–78. https://journals.nnc.kz/jour/article/view/200/199.

President of the Republic of Kazakhstan. *Strategy of the Republic of Kazakhstan on Achieving Carbon Neutrality by 2060*. Decree no. 121, 2 Feb. 2023, unofficial translation. https://unfccc.int/sites/default/files/resource/Carbon_Neutrlaity_Strategy_Kazakhstan_Eng_Oct2024.pdf.

"Refinancing of Barakah Classed as 'Green' Loan." *World Nuclear News*, 13 Mar. 2024. https://world-nuclear-news.org/articles/refinancing-of-barakah-classed-as-green-loan.

Ritchie, Hannah, Pablo Rosado, and Max Roser. "CO_2 Emissions by Fuel." *Our World in Data*, 2020. https://ourworldindata.org/emissions-by-fuel.

Rohde, Robert. "Global Temperature Report for 2025." *Berkeley Earth*, 14 Jan. 2026. https://berkeleyearth.org/global-temperature-report-for-2025/.

Rosatom. "VVER-1200: Flagship Reactor for Kazakhstan." *Rosatom Newsletter*, 25 Aug. 2025, accessed 16 May 2026. https://rosatomnewsletter.com/2025/08/25/vver-1200-russias-flagship-reactor-for-kazakhstan/.

Rosatom. "VVER Reactors: Safe, Reliable, Efficient." *Rosatom Newsletter*, Sept. 2025. https://rosatomnewsletter.com/2025/09/22/vver-reactors-safe-reliable-efficient/.

Samruk-Energy JSC. *Integrated Annual Report 2024: Overview of the Electricity and Coal Market in Kazakhstan*. 2025, pp. 32–43. https://ar2024.samruk-energy.kz/en/kazakhstan-power-and-coal-markets.html.

Satubaldina, Assel. "Kazakhstan's Official Referendum Results Out: 71% Back Nuclear Power Plant Proposal." *The Astana Times*, 8 Oct. 2024. https://astanatimes.com/2024/10/kazakhstans-official-referendum-results-out-71-back-nuclear-power-plant-proposal/.

Sokolova, I. N., N. N. Mikhailova, and A. E. Velikanov. "Induced Seismicity in Areas of Solid Mineral Deposits in Kazakhstan According to the Network KZNET." *Proceedings of the 13th International Conference-School "Problems of Geocosmos"*, 24–27 Mar. 2021, St. Petersburg, pp. 289–99. https://kndc.kz/kndc/docs/publication/2021_2.pdf.

Teollisuuden Voima Oyj. *Nuclear Power Plant Unit Olkiluoto 3*. https://www.tvo.fi/uploads/julkaisut/tiedostot/ydinvoimalayks_OL3_ENG.pdf.

tengrinews.kz. "Earthquake in Karaganda Region: Why Non-Seismic Zones Are Shaking." *Tengrinews.kz*, 1 Nov. 2024. https://en.tengrinews.kz/kazakhstan_news/earthquake-in-karaganda-region-why-non-seismic-zones-are-264832/.

tengrinews.kz. "Who Will Build the Nuclear Power Plant in Kazakhstan?" *Tengrinews.kz*, 16 June 2025. https://en.tengrinews.kz/kazakhstan_news/who-will-build-the-nuclear-power-plant-in-kazakhstan-268296/.

United Nations Economic Commission for Europe. *Energy Policy Brief: Kazakhstan*. UNECE, 2025. https://unece.org/sites/default/files/2025-01/Energy%20Connectivity-Kazakhstan%20Policy%20Brief.pdf.

"VVER-1200: Infographics." *Rosatom Newsletter*, 27 Sept. 2021, accessed 16 May 2026. https://rosatomnewsletter.com/2021/09/27/vver-1200-infographics/.

Xing, Ji, et al. “Advanced PWR Technology—HPR1000 and Unit 5 of Fuqing Nuclear Power Plant.” *Engineering*, vol. 31, no. 12, 2023, pp. 31–36. https://doi.org/10.1016/j.eng.2023.11.009.

Yang, R., J. Wu, G. Gan, and R. Guo. “Revealing the Driving Factors of Water Balance in Lake Balkhash through Integrated Attribution Modeling.” *Rivers and Lakes/Modelling Approaches*, 2025. https://doi.org/10.5194/egusphere-2025-4778.

Deep-Sea Mining: Assessing Environmental Threats and Buoyant Flocculation Rover Solutions By Aaron Zhang, Chris Kim, Noah Kim, Jayden Hwang, Daniel Kim and Ian Oak

Introduction & Background

As land resources begin to diminish, nations are looking to the sea. Currently, deep-sea mining vehicles are sent down to the seabed and use drills and hydro-vacuum technology to extract mineral deposits on the ocean floor (Ashford et al., 2025), typically at 1,000 to 6,000 meters below sea level (Peacock and Ouillon, 2022). The minerals and sediment are then taken up through pipes, and sediment is ejected back into the water column from a mothership at surface level (Ashford et al., 2025). However, the released sediments spread as plumes and have detrimental effects on the ecosystem (Carreiro-Silva et al., 2022). This creates a positive feedback loop where human mining activity destroys benthic (seafloor) ecosystems through poisonous plumes, which spread from 10-100 square kilometers and take tens to thousands of years to settle, depending on the type of sediment (Jeffery, 2020). Research has shown that prolonged exposure to suspended polymetallic sulphide (PMS) particles and Cu^{+} ions released by deep-sea mining poisons cold-water octocorals, killing them in around 27 days. (Carriero-Silva et al., 2022). Coral reefs house around 25% of all biodiversity in the oceans, supporting different species of mollusks, crustaceans, and fish (Vadhel et al., 2025). This presents itself as a striking issue as sea-floor organisms are essential to medicine, such as Marthiapeptide-based cancer treatment (Tortorella et al., 2018) and fisheries that support food security (Williams et al., 2020). These sectors are essential for 70% of the world's countries to further their median household income, education, and life expectancy (Can et al., 2023).

Research Question

To what extent can the impact and spread of sediment plumes be minimized in order to limit the lethal effect on local flora and fauna?

Hypothesis

Through the use of a utilizing hydraulic pumps and flocculation, mining mechanisms can effectively extract ores while ensuring sediments do not create plumes.

Objectives/Aims

The production of sediment plumes has been one of the biggest threats to food security. In order to protect local fauna, a solution must limit or prevent the spread and exposure to suspended particles. It should also allow for deep-sea mining while being beneficial to ecosystems. In creating a solution, we must consider an interdisciplinary approach, incorporating biology, chemistry, and physics to account for all major aspects of life.

Methodology

This research will follow the Design Thinking Framework, outlining the process of empathize, define, ideate, prototype, and test. Firstly, 70% of countries' Human Development Index (HDI), which measures standard of living, education, and life expectancy, has been shown to depend on their respective fishery output (Ali, 2024). Fishery health must be maintained to keep the health and livelihoods of countries. The second step defines sediment plumes as the core problem threatening the safety of cold-water octocoral and associated biodiversity. In order to ideate a solution, we formulated a mechanism combining aspects of physics, chemistry, and biology. A buoyant flocculation rover (BFR) will be attached to the rear of the mining mechanism and use hydrosuction to collect the plumes caused by the initial drilling. Within the BFR, aggregated sediments would be stirred via a motor-powered arm with flocculants in a closed-circuit processing unit and released closely above the seafloor.

Physics

Because the immediate moment of mining –known as the discharge phase– has the least amount of spread, a (BFR) should be in the immediate rear vicinity of the mining vehicle. Flocculants, aggregated masses of sediments, have been shown to settle faster, with flocculant radius (R) and settling velocity (v) being directly proportional ($pf = pw + \frac{9\eta v}{2gR^2}$). The most efficient method to mix flocculants with sediments involves a stirring mechanism (Ali, 2024). Dropping the masses as close to the ground as possible is crucial, since releasing them just 25m above the ground takes ~300 seconds to settle, depending on floc size (Spearman et al., 2020).

Chemistry

The chemical perspective will focus on the formation of floc particles through organic polyelectrolytes. Acting as flocculants, polysaccharides chemically bind small suspended sediments together to form clumps. Molecularly, flocculants are positively charged and will neutralize the particles, which promotes electrostatic attraction and induces floc formation (Aljuboori et al., 2015).

Biology

The aforementioned polyelectrolytes can be found in the form of an organic material known as polysaccharides, which are long chains of carbohydrate monomers. These are optimal for an ocean-based environment due to their efficient flocculating effects and usability across numerous pH conditions. They are found in organic materials such as potatoes, corn, and other starchy plant matter. In addition, bioflocculants are naturally produced polymers by microorganisms (Chen et al., 2017). Therefore, not only are flocculants highly accessible but also very biodegradable, minimizing long-term negative consequences on the ocean ecosystem (Ma et al., 2025). These polysaccharide-based flocculants would be mixed with the collected sediments, creating large masses that easily settle.

During testing, physical analysis will look into the spread and settlement of floc particles. It will focus on the effectiveness of the settling particles. Chemical analysis will record the effectiveness of flocculation in benthic conditions through lab simulations. By simulating deep-sea conditions in a lab environment, the effectiveness of flocculation can be observed. Biological analysis will look into the effectiveness of the bioorganism at producing or serving as a flocculant. It will identify which flocculant is most effective given the situation.

Expected Results/Predictions

Both the short-term and long-term benefits of implementing BFRs would be significant for human society. For one, over 140 countries significantly depend on their fisheries for human development in quality of life and health (Can et al., 2023). Furthermore, deep-sea organisms are imperative for developing antibacterial medicine since they contain compounds such as Marthiapeptide and Desotamide, which are crucial for combating infectious diseases (Tortorella et al., 2018). Therefore, not only will mitigating plume spread allow for further development in life-saving medicine, but it will also allow more than 70% of the world's countries' economic and social developments to continue- developments that could've otherwise been stopped and set back those nations, devastating their people (Source).

Conclusion

As deep-sea mining operations increase in number, there must be countermeasures to the adverse effects of the resulting toxic plumes on benthic life. By implementing BFRs, corals, ocean ecosystems, fisheries, future medicine, and therefore national developments would no longer be in danger.

However, this research includes a couple of limitations. First off, our empathic portion of our methodology relies heavily on statistical data rather than localized testimonials from people in regions where deep-sea mining occurs. Another limitation is that the designed prototype cannot fully replicate deep-sea temperature, pressure, current velocity, biodiversity, and sediment composition due to our finite resources and materials. Our group cannot test which flocculant would be most effective due to our limited resources as well. This weakens the feasibility of our solution since testing is essential to a working prototype. To add on, a final limitation would be the need for legislation to approve our prototype. There are many different rules and regulations that are in place for deep-sea mining machines, established by the National Government as well as the International Seabed Authority (ISA).

Future research will address the current limitations of our research to further improve the feasibility and regulatory viability of our buoyant flocculation rover. To mitigate the overreliance on statistical analysis for the impacts of deep-sea mining on people in regions near where deep-sea mining occurs, researchers should focus on further collaboration with local communities reliant on fishing and the deep-sea. This would allow for a more comprehensive understanding of this issue. Regarding the difficulty of recreating deep-sea conditions, small-scale laboratory experiments that replicate deep-sea conditions could be used to test

flocculants. Additionally, pilot field testing should be utilized to gather real-time data on the behaviors of various flocculants. The results from both experiments could be used to satisfy the third limitation. Both provide scientific, verifiable evidence that can be used to prove that BFR is viable and meets environmental requirements to be given legislative approval by the International Seabed Authority and national governments. Thus, mitigating environmental effects from benthic mining requires a combination of many different aspects of physics, biology, and chemistry. Our research will pave the way for limiting the effects of sediment plumes while also protecting fish food security around the world, allowing countries to prosper economically from deep-sea mining without the adverse effects.

Word Count: 1386

Works Cited

Ali, W. (2024). *Flocculation and deep-sea mining plumes*. Research Repository, Delft University of Technology. https://doi.org/10.4233/uuid:155b3071-c74c-4296-9dd2-512ccb2e1006

Ashford, O., Baines, J., Barbanell, M., & Wang, K. (2025, July 23). *What we know about deep-sea mining—and what we don't*. World Resources Institute. https://www.wri.org/insights/deep-sea-mining-explained

Can, M. F., Mazlum, Y., & Erkan, B. (2023). Is there a relationship between human development and dependence on fisheries? *Oceanological and Hydrobiological Studies, 52*(2), 245–257. https://doi.org/10.26881/oahs-2023.2.09

Carreiro-Silva, M., Martins, I., Riou, V., Raimundo, J., Caetano, M., Bettencourt, R., Rakka, M., Cerqueira, T., Godinho, A., Morato, T., & Colaço, A. (2022). Mechanical and toxicological effects of deep-sea mining sediment plumes on a habitat-forming cold-water octocoral. *Frontiers in Marine Science, 9*. https://doi.org/10.3389/fmars.2022.915650

Coral reef ecosystems. (n.d.). National Oceanic and Atmospheric Administration. https://www.noaa.gov/education/resource-collections/marine-life/coral-reef-ecosystems

Drazen, J. C., DeLong, D., Pakhomov, E. A., Close, H. G., Smith, K. L., Choy, C. A., Clarke, M. E., Jeffreys, R. M., & Sutton, T. (2020). Midwater ecosystems must be considered when evaluating environmental risks of deep-sea mining. *Proceedings of the National Academy of Sciences, 117*(30), 17455–17460. https://doi.org/10.1073/pnas.2011914117

Haffert, L., Haeckel, M., de Stigter, H., & Janssen, F. (2020). Assessing the temporal scale of deep-sea mining impacts on sediment biogeochemistry. *Biogeosciences, 17*(10), 2767–2789. https://doi.org/10.5194/bg-17-2767-2020

Jouffray, J.-B., Blasiak, R., Norström, A. V., Österblom, H., & Nyström, M. (2020). The blue acceleration: The trajectory of human expansion into the ocean. *One Earth, 2*(1), 43–54. https://doi.org/10.1016/j.oneear.2019.12.016

Ma, X., Duan, D., Ji, Y., Du, Z., Shi, L., & Chen, X. (2025). Structural analysis and flocculation performance of ASP-1: A polysaccharide bioflocculant from asparagus root waste. *Industrial Crops and Products, 235*, 121805. https://doi.org/10.1016/j.indcrop.2025.121805

Peacock, T. L., & Ouillon, R. (2022). The fluid mechanics of deep-sea mining. *Annual Review of Fluid Mechanics, 55*(1), 403–430. https://doi.org/10.1146/annurev-fluid-031822-010257

Spearman, J., Taylor, J., Crossouard, N., Cooper, A., Turnbull, M., Manning, A., Lee, M., & Murton, B. (2020). Measurement and modelling of deep-sea sediment plumes and implications for deep-sea mining. *Scientific Reports, 10*(1). https://doi.org/10.1038/s41598-020-61837-y

Tortorella, E., Tedesco, P., Palma Esposito, F., January, G. G., Fani, R., Jaspars, M., & de Pascale, D. (2018). Antibiotics from deep-sea microorganisms: Current discoveries and perspectives. *Marine Drugs, 16*(10). https://doi.org/10.3390/md16100355

Vadhel, H. E., Buhl-Mortensen, L., Mahmoud, M. E. K. M., Vally, Y. E., Meissa, B., Le Loc'h, F., & Menot, L. (2025, July 15). Cold-water coral mounds and reefs off Mauritania and associated megafauna. *Frontiers in Marine Science, 12*. https://doi.org/10.3389/fmars.2025.1575636

Williams, A., Althaus, F., Maguire, K., Green, M., Untiedt, C., Alderslade, P., Clark, M. R., Bax, N., & Schlacher, T. A. (2020). The fate of deep-sea coral reefs on seamounts in a fishery-seascape: What are the impacts, what remains, and what is protected? *Frontiers in Marine Science, 7*. https://doi.org/10.3389/fmars.2020.567002

The Science Behind the Trends: Evaluating Popular Acne Remedies on Social Media Against Clinical Evidence By Aashna Gupta

Abstract: Acne vulgaris is one of the most prevalent dermatological conditions, for approximately 50 million Americans each year, up to 85% of whom are adolescents. Beyond these physical symptoms, acne imposes significant psychosocial burdens, including scarring, anxiety, and depression. Meanwhile, adolescents are among the most active users of social media, with millions engaging with platforms like TikTok and Instagram. Increasingly, they turn to these platforms, not dermatologists, for beauty advice. This has fueled the viral spread of "DIY" acne remedies, which range from citrus-based masks to detox drinks. These are often promoted by uncredentialed influencers. To evaluate the accuracy and safety of these widely circulated remedies, we conducted a two-part investigation: (1) a content analysis of the most-viewed acne treatments on TikTok, Instagram, and YouTube, and (2) a systematic literature review using PubMed and Google Scholar to assess the scientific evidence for each treatment's safety and efficacy. Our findings reveal a clear contrast: the vast majority of popular remedies (such as lemon juice, "skin detox" recipes) lack clinical support and, in many cases, pose risks such as chemical burns or irritation. In contrast, evidence-based treatments such as benzoyl peroxide, topical retinoids, and salicylic acid are consistently underrepresented in social media content despite their proven effectiveness. These results highlight a major gap between dermatological science and social media narratives, with real implications for adolescent skin health. Targeted digital literacy and public health outreach may be necessary to bridge this divide.

Keywords: Acne vulgaris, TikTok skincare trends, Instagram beauty remedies, Remedies, Social Media

Introduction

Acne vulgaris is one of the most prevalent dermatological conditions globally, particularly affecting adolescents. In the United States alone, approximately 50 million people experience acne annually, with up to 85% of teenagers and young adults developing lesions at some point [1]. Beyond its physical manifestations, acne can lead to lasting scarring and profound psychosocial consequences, including anxiety, depression, low self-esteem, and even suicidality, which marks it as a significant public health concern [1].

At the same time, today's adolescents are immersed in digital culture. According to Pew Research, 90% of U.S. teens use YouTube, and over 60% engage regularly with TikTok, Instagram, or Snapchat [2]. Again and again, these platforms serve as their primary source for skincare advice, often in place of medical professionals [3]. While social media offers easy access to information, it also provides a fertile ground for misinformation. Dermatologists have raised concerns about the surge of pseudoscientific content circulating online where acne "cures"

frequently emphasize quick fixes, dietary myths, and unproven home remedies over evidence-based treatments [4,5,6].

Recent studies also confirm this trend: viral content that includes hashtags like #acne or #clearskin commonly promotes DIY approaches such as lemon juice applications or detox drinks [7]. These are often endorsed by influencers without medical credentials [8]. These videos garner millions of views, shaping skincare beliefs and behaviors in ways that are potentially harmful and ineffective [9,10,11].

Literature Review

Our systematic literature review assessed the scientific evidence supporting various acne remedies popularized on social media. The results are summarized in the table below. For Lion Pair Cream, no peer-reviewed clinical trials or systematic reviews were found, categorizing it as having no scientific support, with safety risks unknown beyond general topical cream concerns [12,13]. Tretinoin demonstrated robust evidence from numerous randomized controlled trials (RCTs), confirming its efficacy in reducing both inflammatory and non-inflammatory acne lesions [13]. It is widely recommended in clinical guidelines, though common adverse effects include dryness, peeling, redness, and increased photosensitivity. Benzoyl peroxide and salicylic acid both received strong support from meta-analyses and Phase III clinical trials, confirming their efficacy and safety profiles [14,15]. Benzoyl peroxide has been further validated by recent studies showing no association with cancer risks, although potential safety issues include skin irritation, dryness, fabric bleaching, and rare benzene contamination [16,17]. Salicylic acid, similarly effective, is often used in combination with benzoyl peroxide [18]. PanOxyl, a brand containing benzoyl peroxide, inherits strong scientific backing due to its active ingredient and is frequently recommended by dermatologists despite limited brand-specific studies [19]. In contrast, detox waters and juices, notably those with lemon, lack clinical trial evidence supporting their effectiveness against acne vulgaris. Dermatological sources highlight safety concerns including chemical burns, skin irritation, and increased photosensitivity, categorizing these remedies as unsupported scientifically [20]. Turmeric and honey masks received moderate preliminary support based on small pilot studies and systematic reviews, indicating potential antimicrobial and anti-inflammatory properties [21][22]. However, larger human trials are still needed. Safety issues noted include allergic reactions, skin staining, and variability in dosage and outcomes [22]. Murad Acne Cream lacks independent randomized controlled trials, but its key ingredient, salicylic acid, is supported by moderate evidence. Company-conducted short-term studies suggest a high level of user satisfaction, yet standard safety concerns related to salicylic acid such as dryness and skin irritation apply [23, 24].

Table 1. Relationship between the Social media remedy and scientific evidence for that remedy.

Remedy	**Evidence Strength**	**Safety Notes**
Lion Pair Cream	No Support	Unknown beyond cream concerns
Tretinoin	Strong Support	Dryness, irritation, photosensitivity
Benzoyl Peroxide+ Salicyclic Acid	Strong Support	Irritation, drying, bleaching, rare benzene, contamination.
Panoxyl	Strong Support	Same as Benzoyl peroxide
Detox water/lemon juice	No Support	Chemical burns, photosensitivity, irritation
Turmeric/Honey Mask	Moderate Support	Allergic Reactions, staining
Murad Acne Cream	Moderate Support	Dryness, irritation(salicylic acid effects)

Adapted and created by the authors (Data from Tiktok & Instagram).

Methodology

Social Media Trend Analysis: We conducted a structured content analysis of posts from TikTok, YouTube, Instagram and Reddit. Searches included acne-related keywords (e.g. *"acne cure," "pimple hack," "clear skin hack"*) and popular hashtags (#acne, #clearskin, etc.) which follow methods similar to prior work. We identified the top engagement posts (50-100 per platform) and manually coded each for the key remedy or product featured, influencer type, view/like counts, and claims made (e.g. "X cured my acne overnight"). From this, we built a spreadsheet of the 30-40 most viral remedies, and included the top 7 most viral ones in a table and their purported benefits/safety claims (see excerpt below).

Scientific Literature Review: For each identified remedy, we searched PubMed/Google Scholar using "[*remedy*] AND acne vulgaris." This included specific terms like "tea tree oil," "turmeric AND acne," "benzoyl peroxide AND acne," etc. We prioritized systematic reviews, Cochrane analyses, and RCTs. We extracted data on study design, efficacy outcomes, and adverse effects. Remedies were then classified by the strength of evidence (e.g. *"strong support," "moderate support," "no support"*) and known safety issues.

Data Synthesis: We summarized social-media popularity (views/likes) and scientific support side by side. Numerical metrics from platforms (e.g. TikTok views) were recorded from the content analysis. For evidence grading, we relied on cited trials/guidelines, noting where no trials exist. Safety summaries drew on clinical reports and dermatology resources.

Results

The table below compares popular viral acne remedies with scientific evidence:

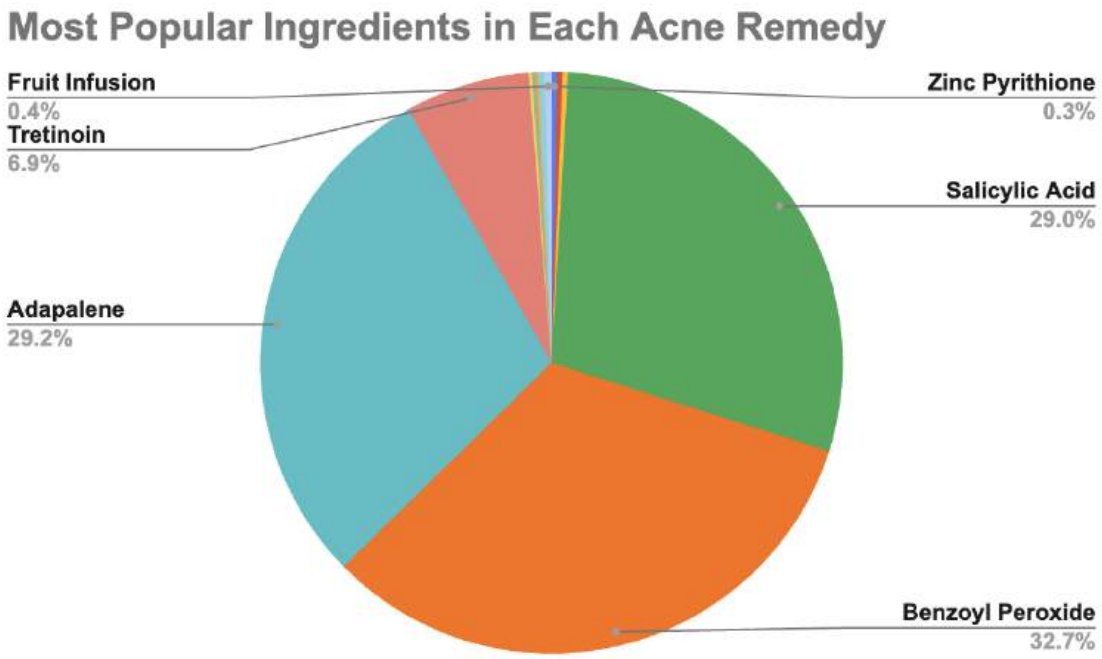
Table 2. Correlation between Popular Viral Acne Remedies and the Scientific Evidence

Remedy	**Social Media Trend (Approx)**	**Claimed Effects**	**Scientific Evidence**	**Safety Concerns**
Lion Pair Cream	308K+ likes on TikTok	Acne is gone; Acne scars are gone	Limited or mixed evidence; varies by ingredient	Unknown beyond cream concerns; one should seek a dermatologist before using it
Tretinoin	417K+ likes on tiktok	Reduces inflammation and wrinkles	Strong clinical support. Prescription retinoid that increases cell turnover and reduces acne	High potential for dryness, photosensitivity, and irritation.
Benzoyl peroxide, and Salicylic Acid	546K+ likes on Tiktok	Reduces acne	Strong support. FDA approved acne treatment shown to reduce acne causing bacteria and inflammation.	Can cause dryness, irritation, and bleaching of fabrics.
Panoxyl	211K+ likes on Tiktok	Skin Glows	Contains Benzoyl Peroxide which is strong evidence for acne reduction.	May cause irritation or dryness to the eyes; avoid contact with eyes.
“Detox” water /	22,170 likes on Instagram	“Detoxifies skin	No evidence. Drinking water or juices has no known effect on acne	No evidence. Drinking water or

juices (e.g. lemon water)		internally, prevents acne”	clearance. These claims stem from general health myths. There are no trials testing “detox” beverages for acne.	juices has no known effect on acne clearance. These claims stem from general health myths. There are no trials testing “detox” beverages for acne. Harmless if non-citrus, but citrus juices still risk skin irritation if applied topically. Ingesting excessive citrus can upset the stomach.
Turmeric / honey mask or patch	10,639 likes on Instagram	“Anti-inflammatory; reduces redness/bumps”	Limited preliminary evidence. Curcumin (turmeric’s active) has shown antimicrobial and anti-inflammatory effects in lab and animal studies. Some small trials hint it may help inflammatory acne, but no robust human RCT confirms efficacy	Generally safe to eat; topical turmeric can stain skin and sometimes cause contact dermatitis. Honey may occasionally irritate acne-prone skin. High-dose supplements can cause GI upset.
Murad Acne Cream	2K+ likes	Makes acne bumps flatter	Limited evidence.	May cause dryness to the skin. Seek a dermatologist.

Adapted and created by the authors (Data from Tiktok & Instagram)

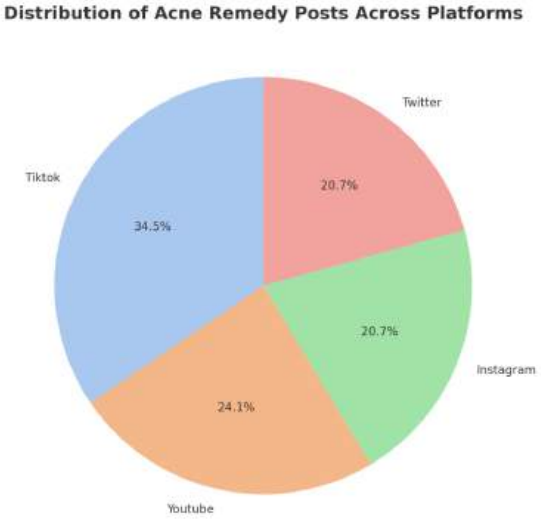
Figure 1. Most Popular Ingredients in Each Acne Remedy



Adapted and created by the authors (Data from Tiktok & Instagram).

The pie chart above further reveals that the majority of viral remedies in our dataset feature dermatologist-approved ingredients. Benzoyl peroxide appears most frequently (32.7%), followed closely by adapalene (29.2%) and salicylic acid (29.0%). These ingredients are widely supported by clinical research and are recommended in most acne treatment guidelines, indicating that while social media content often lacks context or professional oversight, many users still gravitate toward products with scientifically validated ingredients. In contrast, ingredients like tretinoin (6.9%),despite strong evidence, appear less frequently, possibly due to their prescription-only status. Meanwhile, ingredients such as zinc pyrithione (0.3%) and fruit infusions (0.4%) are rare and reflect more experimental or anecdotal approaches.

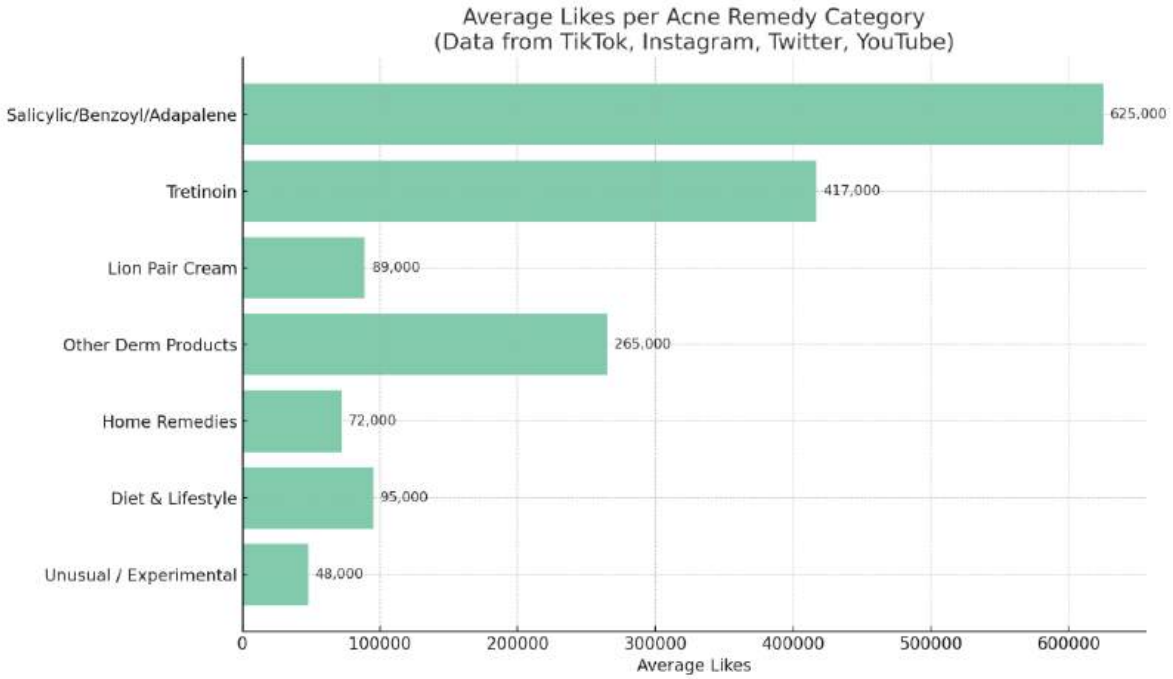
Figure 2. Distribution of Acne Remedy Posts Across Platforms



Adapted and created by the authors (Data from Tiktok & Instagram).

As seen in the Pie Chart in *Figure 2*, TikTok accounts for the largest share of viral acne remedy posts (34.5%), followed by YouTube (24.1%), while Instagram and Twitter each contribute 20.7%. This distribution reflects how platforms known for short-form or viral content, especially TikTok, play an outsized role in shaping skincare narratives online. Our research confirms that many TikTok posts promote trendy but unverified remedies, such as Lion Pair Cream or DIY acne routines, often accompanied by dramatic before-and-after claims. YouTube, while more balanced, still hosts a mix of credible dermatology advice and anecdotal or influencer-driven solutions. Instagram posts frequently center on holistic or aesthetic remedies like turmeric masks or detox waters, which tend to lack strong clinical backing. Twitter posts, being shorter and less detailed, often spread simplified advice without scientific context. The chart reinforces the key concern in our research of how the most popular platforms for acne remedy content are also those where scientific evidence is often weakest, thus why misinformation can spread rapidly.

Figure 3. Average Likes Per Acne Remedy Category



Adapted and created by the authors (Data from Tiktok & Instagram).

This bar graph shows the average number of likes each acne remedy category received across TikTok, Instagram, Twitter, and YouTube. Clinically proven treatments like salicylic acid, benzoyl peroxide, and adapalene had the highest engagement (~625K likes), followed by tretinoin and dermatologist-recommended products such as PanOxyl and CeraVe. Interestingly, diet and lifestyle changes and home remedies (such as turmeric, lemon) also gained considerable attention, despite weaker scientific support. Lion Pair Cream, a popular Japanese product, received moderate interest. Meanwhile, experimental methods like anti-dandruff shampoo or topical insulin had lower, but still notable, engagement. Overall, the data suggests that social media users are drawn to both science-backed treatments and trending natural or unconventional solutions. Our analysis illustrates a clear contrast between the acne remedies most widely circulated on social media and those supported by dermatological research. Across platforms (particularly TikTok and Instagram) we found that the majority of viral acne treatments lacked strong scientific validation. Although medically endorsed treatments such as topical retinoids (e.g., tretinoin), benzoyl peroxide, and salicylic acid were present in the dataset, they were relatively underrepresented and often embedded within broader, less substantiated skincare routines. For example, a post might mention salicylic acid, but center the routine on more attention grabbing and non-clinical elements like ice facials, turmeric-honey masks, or "detox" drinks. Among the most frequently viewed remedies were Lion Pair Cream, tretinoin, salicylic acid, lemon face mask, and panoxyl. Many of these, particularly Lion Pair Cream and lemon based treatments, were praised for immediate results, yet lacked strong scientific and clinical

evidence and posed safety concerns such as skin irritation and acid burns. In contrast, dermatologist-approved options like tretinoin and salicylic acid, while backed by strong scientific support, often received less engagement. Based on this disparity, we were able to conclude that social media algorithms and influencer narratives prioritize novelty and virality over evidence-based care, and that the popularity of a certain remedy does not correlate with its medical credibility, this case being acne treatment. The quantitative analysis of our dataset suggests that approximately 70% of the top viewed remedies lacked substantial clinical support.

Furthermore, roughly 35-40% posed potential harm, particularly when used incorrectly or in combination with other harsh substances. Some documented risks included chemical burns from lemon juice, excessive dryness or peeling from retinoids, and contact dermatitis from turmeric-based masks. In rare but concerning cases, we found anecdotal evidence of users experiencing worsened breakouts, long-term skin discoloration, or allergic reactions after following influencer-endorsed routines without professional guidance. The discrepancy between popularity and scientific reliability raises serious public health concerns.

Discussion

Adolescents and young adults who are already vulnerable to the psychosocial impacts of acne may be misled by these viral trends, delaying or entirely avoiding treatments that are both safe and effective. In some cases this misinformation could lead to worse outcomes, including prolonged acne flares, increased risk of scarring, or new injuries from inappropriate topical applications. Moreover, the normalization of untested remedies on social media contributes to widespread confusion about how acne actually develops and what treatments are supported by medical science. Dermatologists have warned that the credibility of health content on social platforms varies widely depending on the source. Posts that are created by board certified dermatologists or other licensed professionals often provide accurate and balanced information, yet these are outnumbered by content from influencers with no formal training.

Our findings support these concerns: in our dataset, posts from non-medical creators were not only more common, but also more likely to include exaggerated claims such as “clears your skin overnight” or “cures acne naturally in days”. Multiple peer reviewed studies on health misinformation online confirm that adolescents are especially susceptible to such claims, particularly when they are visually appealing and endorsed by attractive influencers. This points to a broader issue in digital health literacy. While social media has the potential to democratize health education, it also risks amplifying pseudoscience when left unmoderated. Our results reinforce the need for stronger evidence based communication on platforms heavily used by youth. Some interventions such as promoting verified dermatologists on TikTok or including disclaimers on unverified medical claims may help reduce harm. Narrowing the gap between dermatological science and social media skincare culture will require collaboration between medical professionals, educators, public health organizations, and the platforms themselves.

Acknowledgements

We would like to thank our parents for supporting us while writing this.

Limitations

Our study has several important limitations. Firstly, the content analysis focused only on the most popular posts across selected social media platforms, which may not fully represent the breadth of available content or capture less visible but potentially influential posts. Secondly, popularity metrics (views, likes) were more likely than not influenced by factors unrelated to the perceived effectiveness or safety of remedies, such as influencer popularity or platform algorithms, potentially skewing the data. Thirdly, the literature review was limited to articles available in PubMed and Google Scholar databases, possibly overlooking relevant studies published in non-indexed journals or in languages other than English.

Conclusion

Our analysis reveals a significant difference between the most viral acne remedies on social media and those supported by clinical dermatology. Approximately 70% of the top viewed treatments lacked substantial scientific backing, and over one third posed potential safety risks such as skin irritation, acid burns, or allergic reactions, often when misused or combined with other harsh ingredients. Remedies such as lemon face masks and Lion Pair Cream were widely promoted despite little to no peer-reviewed evidence. Meanwhile clinically endorsed treatments such as benzoyl peroxide, carbon facials, and salicylic acid, while highly effective, were relatively underrepresented and often overshadowed by more sensational and influencer-driven routines.

This gap sparks the urgent need for digital health literacy and intervention: social media platforms could prioritize vetted content, dermatologists might expand their online presence, and educators and families can help adolescents critically evaluate skincare claims. Increasingly, teens should be equipped with accurate, evidence based guidance to protect their skin health and reduce the harm caused by viral misinformation.

References:

1. American Academy of Dermatology. (n.d.). *Acne: Diagnosis and treatment*. https://www.aad.org/public/diseases/acne/derm-treat/treat
2. American Academy of Dermatology. (n.d.). *Skin conditions by the numbers*. https://www.aad.org/media/stats-numbers
3. Atkin, N. J. (2020). From high-flying trapeze to highly infectious disease: The history of dermatology in the circus. *Clinical Dermatology*, *38*(5), 654-660. https://doi.org/10.1016/j.clindermatol.2020.09.006
4. Adebesin, F., Smuts, H., Mawela, T., Maramba, G., & Hattingh, M. (2023). The role of social media in health misinformation and disinformation during the COVID-19 pandemic: Bibliometric analysis. *JMIR Infodemiology*, *3*(1), Article e48620. https://doi.org/10.2196/48620
5. Bal, Z. I., Karaosmanoglu, N., Temel, B., Bas, V. M., Karali, M. G., Ergun, E. Z., Avci, E. B., Tellal, E. S., Yuksel, E. I., & Nalbant, S. (2025). Trust in dermatologists versus social media influencers among acne patients. *Cureus*. https://doi.org/10.7759/cureus.57021
6. Dréno, B., Bettoli, V., Araviiskaia, E., Ochsendorf, F., & Layton, A. (2018). The influence of exposome on acne. *Journal of the European Academy of Dermatology and Venereology*, *32*(9), 1445-1455. https://doi.org/10.1111/jdv.14820
7. Ertekin, S. S., Salici, N. S., Manav Bas, V., Karali, M. G., Ergün, E. Z., Avcı, E. B., Tellal, E. S., & Yüksel, E. I. (2024, July). Influence of social media and internet on treatment decisions in adult female acne patients: A cross-sectional survey study. *Dermatology Practical & Conceptual*, *14*(3), Article e2024156. https://doi.org/10.5826/dpc.1403a156
8. George Washington University. (2023). *Study finds TikTok filled with acne misinformation*. https://mediarelations.gwu.edu/study-finds-tiktok-filled-acne-misinformation
9. Hughes, M. M., Shaw, K. A., Patrick, M. E., & DiRienzo, M. (2023). Adolescents with autism spectrum disorder: Diagnostic patterns, co-occurring conditions, and transition planning. *Pediatrics*, *152*(3), Article e2022060085. https://doi.org/10.1542/peds.2022-060085
10. Hunt, I. de V., & Linos, E. (2022). Social media for public health: Framework for social media-based public health campaigns. *Journal of Medical Internet Research*, *24*(12), Article e42179. https://doi.org/10.2196/42179
11. Leyden, J., Stein-Gold, L., & Weiss, J. (2017). Why topical retinoids are mainstay of therapy for acne. *Dermatology and Therapy*, *7*(3), 293-304. https://doi.org/10.1007/s13555-017-0185-2
12. Liu, Y., Dan, Y., Yang, J., He, X., Liu, J., Yi, Y., Chen, X., Yin, X., Song, W., Niu, Y., Zheng, Y., & Ai, Y. (2025). Clinical efficacy of a salicylic acid–containing gel on acne

management and skin barrier function: A 21‐day prospective study. *Journal of Cosmetic Dermatology*, *24*(7), Article e70353. https://doi.org/10.1111/jocd.70353
13. Medical News Today. (n.d.). *Turmeric face masks*. https://www.medicalnewstoday.com/articles/turmeric-face-mask
14. Medical News Today. (2024, June 10). *Everything you need to know about cystic acne*. https://www.medicalnewstoday.com/articles/103258#prevention
15. Montgomery, K. C., & Chester, J. (2001). Youth and digital media: A policy research agenda. *Journal of Broadcasting & Electronic Media*, *45*(1), 138-150. https://doi.org/10.1016/s1054-139x(00)00130-0
16. O'Connor, C., O'Grady, C., & Murphy, M. (2024). Spotting fake news: A qualitative review of misinformation and conspiracy theories in acne vulgaris. *Clinical and Experimental Dermatology*, *49*(1), 1-8. https://doi.org/10.1111/ced.15222
17. Ogwu, M. C., & Izah, S. C. (2025, March 1). Honey as a natural antimicrobial. *Antibiotics*, *14*(3), Article 255. https://doi.org/10.3390/antibiotics14030255
18. Pappas, A. (2009). The relationship of diet and acne. *Dermato-Endocrinology*, *1*(5), 262-267. https://doi.org/10.4161/derm.1.5.10192
19. Pew Research Center. (2023). *Teens, social media and technology 2023*. https://www.pewresearch.org/internet/fact-sheet/teens-and-social-media-fact-sheet/
20. The Dermatology Digest. (n.d.). *PanOxyl product recommendations*. https://thedermdigest.com/survey-panoxyl-named-no-1-dermatologist-recommended-otc-benzoyl-peroxide-brand/
21. Wojtara, M. S. (2023). Use of social media for patient education in dermatology: Narrative review. *JMIR Dermatology*, *6*, Article e42609. https://doi.org/10.2196/42609
22. Yousaf, A., Hagen, R., Delaney, E., et al. (2021). The influence of social media on acne treatment: A cross-sectional survey. *Journal of the American Academy of Dermatology*. https://doi.org/10.1111/pde.14091
23. Zaenglein, A. L., Pathy, A. L., Schlosser, B. J., Alikhan, A., Baldwin, H. E., Berson, D. S., Bowe, W. P., Chien, A. L., Cohen, J. L., Collyer, A. J., Del Rosso, J. Q., Draelos, Z. D., Gold, L. F., Granger, C., Harper, J. C., Keri, J. E., Kircik, L. H., Levy, M. L., Leyden, J. J.,... Bhushan, R. (2016). Guidelines of care for the management of acne vulgaris. *Journal of the American Academy of Dermatology*, *74*(5), 945-973. https://doi.org/10.1016/j.jaad.2006.08.048
24. Zhang, Z., & Zhou, M. (2024, August 1). The impact of social media information exposure on appearance anxiety in young acne patients: A moderated chain mediation model. *Frontiers in Psychology*, *15*, Article 1409980. https://doi.org/10.3389/fpsyg.2024.1409980

Why Nations Fail Review By Yağmur Eren

Why Nations Fail? was written by Daron Acemoglu and James A. Robinson in 2012. Since the day it was published, the book has caused sensations in the fields of economics and politics. Daron Acemoglu won a Nobel Prize in the field of economics in 2024 and is one of the main architects of literature in the field of institutional economics. Robinson is a professor of politics and political economy at the University of Chicago. Robinson's research focuses on the effects of historical and political institutions on the economic development of countries. These academics collaborated extensively, and their works include *The Narrow Corridor* which was published in 2019. Their collaborative works usually focus on the effects of political institutions on economic development and growth.

Why Nations Fail? posits that the reason some countries are poor and some are wealthy depends on the structure of their political institutions rather than geographical, cultural, or religious reasons. This approach makes the book stand out among its rivals because they address the reasons of the problems clearly and to the point. Acemoglu and Robinson suggest that inclusive political institutions protect property rights in order to incentivize individuals to invest. In contrast, power corrupted extractive political institutions collect the resources held by the elite and reallocate them to the citizens. By employing a variety of case studies, this book offers fresh insights into the subject. Moreover, the book is impressive for its interdisciplinary approach. It is not only a successful text on economics but also encompasses history, politics, and sociology. It is wide-ranging geographically too. The book uses the examples of North and South Korea, Nogales, Latin America, etc., which helps to show how some countries in the same geographical area, with similar cultures but different institutions, have very different economic trajectories. Ultimately though, as the book aims to demonstrate why some countries are wealthy and some are poor, it is first and foremost an economics text.

According to the book, economic development in a country is dependent on its political institutions. Inclusive institutions distribute political power among a wide group of people and also protect the property rights of individuals. These conditions make individuals more confident about making investments, which also means that these policies incentivize people to invest and take risks. Thus, economic growth is achieved. On the other hand, extractive institutions are the opposite of inclusive institutions. Extractive institutions have a limited political perspective which is shaped by the society's elite. Accordingly, these countries exploit their resources to satisfy the needs of this elite group. In extractive institutions, social mobility is restricted and research and development is limited. Therefore, they are unlikely to become prosperous (except for the ruling elite).

The book emphasizes that the economic trajectory of countries is shaped by the decisions

they make at critical junctures. In order to illustrate this argument, the book presents detailed case studies. These studies, from diverse countries, also help to define the phrase “critical junctures.” One of the first case studies presented in the book is the town of Nogales being split into two. Separated only by a border line, one side is in Arizona, USA and the other is in Sonora, Mexico. Despite their shared geography, culture, language, and even having a similar demographic ancestry, there is a huge disparity in living standards between each side of the border. The major difference between them is in their political institutions. On the American side of Nogales, new projects are incentivized and living standards are high. On the Mexican side, economic activities are limited, power is in the hands of elite and the government is oppressive. Also by comparing pairs like the UK and Spain, North America and Latin America, North Korea and South Korea, Botswana and Congo, and China and Spain, the book creates a structure of contrasts and opposition for the reader, which strengthens their thesis.

The book consists of fifteen chapters. In the first chapter, through the example of Nogales, the difference in living standards is explained. In the second chapter it is posited that geography, culture, and ignorance are not only reasons for poverty. The third chapter underlines how inclusive and extractive institutions were born. In the fourth chapter, significant turning points in history are explained. In the fifth chapter, through the example of the USSR, the limitations of centered planning are exposed. The sixth chapter explains how different areas in the world have followed different trajectories. The seventh chapter discusses historical revolutions like the Glorious Revolution (1688) in the UK. Chapter eight demonstrates how the failure of the elite group to embrace technological development and improvements limits progress and prosperity. The ninth chapter examines the causes of underdevelopment in certain countries. The tenth chapter discusses wealth distribution and how this affects countries’ economies. The eleventh chapter reveals how inclusive institutions strengthen their country and how this leads to a positive ‘feedback’ cycle. Chapter twelve demonstrates how extractive institutions create poverty by their decisions. The thirteenth chapter explains how nowadays political institutions fail to achieve economic growth and leave the poverty trap. The fourteenth chapter suggests ways in which some countries could break this cycle. In the last chapter the book goes over the significance of the role of political institutions for a country’s growth and summarizes the theoretical frame of the thesis. Throughout these fifteen chapters, the authors cite historical evidence and offer historical comparisons in order to support their arguments and thesis. Notably, the book covers a substantial time period. Also, in order to make the important keywords easier to understand, authors explain them with detailed examples. For instance, through examples like the UK, France, the Netherlands, etc.

The main argument presented in the book is that institutional corruption affects the economic system of countries. In other words, the decisions made by political institutions cause a country to be wealthy or poor. The book refers to two types of institutions: extractive and

inclusive institutions. Inclusive institutions distribute the power they have into all groups in a society; it is not only in the hands of the elite. Thus, other sectors of society have influence on the way the country is governed. This has positive effects on the economy as well. In inclusive institutions individuals have access to loans, individuals can invest, and competition is present. The other type of political institution is what is termed an extractive institution. In these kinds of institutions, political power is limited, and it is in the hands of the elite. These kinds of institutions don't work for the sake of the whole society; rather, they focus on satisfying the needs of the elite. Consequently, these systems have negative effects on the economy. In extractive institutions resources are exploited, opportunities are limited, and innovation is not incentivized; in fact, it is oppressed. In these kinds of institutions, consumers assume that their accumulated wealth will be taken from them, so they become risk-averse.

Another key point mentioned in the book is the importance of incentivizing people to innovate and produce. The book highlights that in countries that protect property rights, where the government's decisions are consistent and the justice system works correctly, people are more educated, are risk-takers (because they feel able to take risks), and have long-term financial/investment plans. These institutions epitomize the meaning of inclusive institutions. Furthermore, the book also demonstrates that in some countries, extractive institutions exploit their power and often operate within conflicts of interest. Overall the book states that institutions, technological development, and political constraints have a significant role in shaping the future of nations.

Even though the book has strong arguments and it is an engaging and well-researched work, it does not present a perfect theoretical framework. In order to support their arguments, the authors use examples like the UK, France, the Netherlands, etc., to support the idea that European countries are economically strong. However, when the theory is applied to some countries in the past and some that still go on, such as India, the Ottoman Empire, and Venice, the limitations of this theory can be observed. For instance, the printing press was not allowed in the Ottoman Empire. This means the book explains Western Europe's century's worth of experience in a detailed way, but the framework they explain it in cannot be applied to other countries, which is a limitation of the theory.

Additionally, in some chapters, geographical and other factors are not accounted for and everything is viewed from a political and economic perspective. For example, America's discovery, the discovery of new trade paths, and some regions having low population due to illness are not considered causes of economic conditions. Also, the authors' use of the word "inclusive" is vague and generalized. The same word is used to describe both ancient and modern democracies, which confuses the meaning and devalues the word. Conversely, other key words are defined in detail. However, the book considers conflicting ideas without prejudice, and they

present their arguments as open to interpretation and discussion. This honesty, combined with its interdisciplinary approach, makes the book more powerful and its arguments more convincing.

Why Nations Fail? is a book that should be a bedside book for economics and politics students. At the same time, researchers and academics should also view this book as a valuable resource. The book is an excellent way to understand the global economy from a historical perspective, which illuminates our understanding of economic conditions and their causes. I believe Daron Acemoglu and James A. Robinson's book is a seminal work in this field and will continue to influence academics, students, and anyone who is interested in the global economy for some time to come.