|CURIEUX|

ACADEMIC REVIEW

Humanities & Social Sciences

2026 Issue Part 3

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: WHEN VILLAGES VANISHED An Examination of Nicolae Ceaușescu's Regime and Its Effects on Romanian Cultural Identity By Alexander Dobrescu

Page 19: Conscience and Constraint: Lillian Hellman's Testimony and the Limits of Cold War Resistance By Evelyn Cai

Page 29: Global Energy and International Relations: A Study of Energy Security, Geopolitics and the Global Energy Transition By Toshani Padarthi

Page 51: Code-Switching in Shanghai International Schools: Contexts, Motivations, Effects, and Implications By Pete Chen

Page 61: The Bilingual Brain: Does Language Dominance Affect Cognitive Flexibility? By Ziyu Wang

Page 81: Foreign Policy in Real Life: How Teenagers Perceive the Relevance of Global Issues to Their Daily Lives By Emma Orford

Page 94: Virtue or Verdict: A comparative study of Sulh-i-Kul and Bhinneka Tunggal Ika on Legal Pluralism, Sovereignty and Historical Legacy By Prathama Mandal

Page 100: Why Habits Override Conscious Decisions? By Aanya Sodhi

Page 107: Who Do We Trust? Attachment, Anthropomorphism, and Age in Human–AI Relations By Ahana Gandhi

Page 124: American Youth Desensitization to Political Violence By Leighton Westerberg

Page 145: Two Generations, One Amendment: Elizabeth Cady Stanton, Alice Paul, and the Era-Dependent Tactics of the American Woman Suffrage Movement By Sarina Singhal

Page 155: The Correlation of the Normalization of Mistreatment in Reproductive Health Systems and The Yellow Wallpaper By lillian korzonkiewicz

Page 160: The Flaws of the Indian Reorganization Act of 1934 By Solomon Heyman

Page 166: Precarity or Belonging? Job Insecurity, Social Capital, and Well-Being among Immigrant Workers in the Chinese Ethnic Enclave Economy By Elaine Liu

Page 181: The 'Last Mile' of Skilling: How Social Enterprises Can Bridge the Gap Between Vocational Training and Actual Employment for People with Disabilities By Ayaan Jain

Page 188: Mapping Access to Urban Green Infrastructure in Bengaluru By Sanya Pai

Page 196: Exploring the Role of Dual N-Back Training in an Academic Setting Among the Pre-Adolescent Population By Tamanna Sethia

Page 206: Can Poetry Cultivate Empathy? A Review of Psychological Mechanisms and Empirical Evidence By Penny Wei

Page 221: The significance of George Balanchine's *Agon:* A groundbreaking neoclassical creation of push and pull By Sonia Schultes

Page 231: The Untold Heroes Behind the Line in the Second World War By Daniel Meng

Page 241: Impact Of Family And Peers On Social Identity Formation In Adolescence By Aanya Ghosh

Page 251: Evaluating the Efficacy of a Low-Cost Digital Learning Platform on Academic Performance of Students Aged 5-16 By Ansh Tripathi

Page 266: Micro-Activism, Algorithms, and Youth Civic Engagement in the Age of Short-Form Video: A Review of the Literature By Trisha Mittal

Page 286: Artificial Intelligence Personalization and the Changing Social Dynamics of Sports Fandom By Simon Zhong

Page 298: How Does Patent and Copyright Law Affect Innovation in the Video Games Industry By Almira Seissembayeva

Page 305: To What Extent Can Behaviours Commonly Associated with Psychopathy Be Explained by Social Roles, Socialization, and Institutional Environments? By Andy Peng

Page 316: Coaches, Critiques, and Confidence: A Mixed-Methods Study on Body Image in Retired Competitive Gymnasts By Bradie Hinson

Page 332: Federal Architecture and the Institutionalization of Racial Risk: New Deal Housing Policy, Mortgage Finance, and the Durable Structure of Inequality (1933–1954) By Aarav Dhar

Page 354: The Inherent Mathematical and Philosophical Natures of Counterintuitive Decision-Making Within Society By Christopher Kim

Page 368: Greenwashing: A Review of the Linguistic Strategies Corporate Use By Lauren Kim

Page 374: Images of Violence: An Analysis of Cartel Imagery in Media and Its Impact on Mexican Society By Alston Guo

Page 380: Where the Music Takes You: Autobiographical Recall vs. Fictional Narrative Across Classical Subgenres By Derrick Lin

Page 397: How do Social Media Platforms Transform Fandom Rituals Into Modern Forms of Collective Practice? By Siyan Wang

Page 415: An Archive of Our Own: How Fans Seized the Means of (Re)Production By Ruisi He

Page 434: A Cruel Justice: *The Justice of Southern Communities and Norms In Mark Twain's Novel The Adventure of Tom Sawyer* By Thomas Rutterer

Page 440: Memory, Belonging, and the Gradual Normalization of Exclusion: A Case Study of Holocaust Survivor Gabriella Karin By Honor Chacon

Page 452: Procedural and Conceptual Knowledge: a Chronological Relation in Basic Derivatives By Quinn Harman

Page 462: The Evidence of the Dead: Material Truth and Political Memory in Lucan's *Pharsalia* By Derek Hu

Page 470: Factors of Happiness in Various Cultures By Melia Outman

Page 484: LA County Medical Professional Perspectives on Physician-Assisted Suicide in Relation to Ethical Theory and International Law By Noa Sarfati

Page 506: Plural Emotion Readings in Short Social Video Clips: A Descriptive Study By Evelyn Kim

Page 525: Shell Shock Post World War I: Challenging and Encountering Masculinity By Talia Fine

Page 532: How Do Historical Medical Injustices Affect Trust in Gene Therapy Among Minority Communities? By Kayla Nicole Ofosu

WHEN VILLAGES VANISHED
An Examination of Nicolae Ceaușescu's Regime and Its Effects on Romanian Cultural Identity By Alexander Dobrescu

Abstract

From 1965 to 1989, Nicolae Ceaușescu tried to reshape Romania around his own vision of a “modern” state. His policies went far beyond politics and changed many parts of everyday Romanian life. The Securitate created widespread fear through surveillance and intimidation, making open expression dangerous and weakening trust between people. At the same time, Ceaușescu’s economic policies left many Romanians without enough food, heat, or electricity and forced daily life to revolve around basic survival. Systematization destroyed villages, churches, and historic neighborhoods, while propaganda reshaped Romanian history around Ceaușescu and his regime. These policies weakened many of the communities, traditions, and social bonds that had helped shape Romanian identity. Evidence from the Letter of the Six, Ceaușescu’s 1989 trial, diplomatic accounts, and historians of communist Romania shows that the regime created obedience through fear, but never built lasting loyalty. This weakening of Romanian society contributed to the anger that eventually erupted in the 1989 Revolution.

Introduction

"Down with Ceaușescu!" roared the crowd in Brașov in November 1987, two years before the protests in Timișoara sparked the start of the Romanian Revolution (Petrescu and Petrescu 333). After twenty-four long years of Nicolae Ceaușescu's rule, the Romanian people finally rose up against him. From 1965 to 1989, Ceaușescu transformed Romania from a semi-liberal, yet still authoritarian, nation to the Eastern Bloc's most brutal, suppressive communist state. He censored the Romanian people, destroyed neighborhoods, and plunged the country into economic ruin, all while limiting the Romanian people's rights. Ceaușescu's Securitate, the nation's secret police force, acted as the main social enforcer of his policies, suppressing freedoms and dissent, functionally eroding the communal bonds of the Romanian people. By the end of the 1980s, Romania stood isolated and daily life turned miserable. Most households lacked heat, food, and electricity, vital necessities in Romania's bitter winters that often fell below freezing temperatures. Pushed to the brink, the Romanian people began fighting back and a revolution erupted in December of 1989, leading to the collapse of the regime and Ceaușescu's summary trial and execution. Historians often debate why certain authoritarian regimes collapse whereas others persist. Romania's sociopolitical and economic conditions ultimately resulted in the former, a cautionary tale showing that a leader who uses force can breed greater resentment. Ceaușescu's imposition of his counterfeit vision of a "modern" Romania, enforced through oppression, tore the nation away from traditional sources of Romanian identity, eroding Romanian society and contributing to a revolution. The political oppression from the Securitate, the regime's grave economic mismanagement, and the

destruction of culture through systematization and propaganda all heavily contributed to a sharp departure from Romanian identity, eroding the values integral to tradition.

Village Life, Orthodoxy, and Latin Heritage Shaped a Distinct Romanian Identity

Romania possesses a rich cultural history, developed over a period of thousands of years. Traditionally, social life was rooted in Romanian villages, which served as the center of Romanian culture as families shared resources, preserved oral traditions through folk songs, and passed down stories through generations. As historian Henri H. Stahl documented, peasant communities preserved traditions from the thirteenth century through communal villages known as obștea, where extended families cultivated ancestral lands, shared resources, and transmitted oral traditions across generations (Stahl). The Romanian Orthodox Church also provided and cultivated a deep cultural connection. The Orthodox faith permeated daily life in many aspects, from marriages and deaths to the calendar, Easter celebrations, and saint day celebrations, all of which bound communities together. Anthropologist Gail Kligman's fieldwork in Maramureș revealed how these life cycle rituals, especially mourning ceremonies, sustained these social bonds (Kligman 211–14). The Romanian language also connected the nation to its Latin heritage. Romania is a country surrounded by Slavic neighbors, so Romanian carried even more weight as a Latin language. The language has historical memory embedded in its vocabulary and expressions, all with concepts like *cinste* (honor and integrity) rooted in the cultural identity of Romania. These three elements---village culture, Orthodox tradition, and the Romanian language---formed part of the Romanian identity that Ceaușescu would systematically target and erase.

To understand the conditions that led to the disruption of Romanian identity, a brief history outlining Romania's evolution as a nation is necessary. Romania, often known as a "Latin island in a Slavic sea," traces its linguistic and cultural heritage to Roman colonization of Dacia in 106 CE under Emperor Trajan (Bachman 3). The Romans introduced Latin to the region, whereupon the people adopted it as their everyday language, and even though Slavic people surrounded the area, Romanian still kept its Romance language structure. The Romanian language helped preserve a cultural link to Western Europe and, with it, a distinct Romanian identity. This identity was also shaped by Romanian Orthodoxy. After the Great Schism of 1054, Romanian people began to follow Eastern Orthodoxy. Romania developed its own form of Orthodoxy, shaped by its own character while blending Byzantine religious practices and cultural aspects. After Constantinople fell in 1453, the church played an even larger role in daily life as the churches helped preserve the language, traditions and beliefs of the people.

The Romanian people underwent centuries of vassalization under the rule of the Ottoman Empire. The Ottomans imposed tribute payments on Wallachia and Moldavia and restricted political autonomy. However, unlike other territories under Ottoman rule, the Romanian people maintained some form of independent governance and crucially religious freedom to practice their Orthodox faith. The religious freedom allowed the Romanian Orthodox Church to help

preserve the Romanian language and traditions through churches, monasteries, and village parishes.

The unification of Wallachia and Moldavia on January 24, 1859 under Alexandru Ioan Cuza, the eventual prince of the newly formed Romanian Principalities, laid the groundwork for this independence (Bachman 26). Later on, facing mounting pressure from the Romanian elites, Cuza abdicated, and Carol I of the Hohenzollern-Sigmaringen dynasty was installed in 1866 beginning a new era of constitutional monarchy. Following the Russo-Turkish War of 1877–1878 Romania gained full independence from the Ottoman Empire, drawing the country closer to Western European norms (Bachman 28). During the interwar years Romania also attempted to create a democracy, but ethnic issues, political violence, and economic turmoil undermined stability in Romania. The Great Depression only further exacerbated these issues.

Romania's trajectory towards communist rule began with catastrophic decision making during World War II. In 1940, General Ion Antonescu seized power through a military coup alongside the fascist Iron Guard movement, overthrowing the government and installing a dictatorship. Antonescu allied Romania with Nazi Germany which resulted in the deportation and murder of hundreds of thousands of Jews and Roma (International Commission). However, Romania's trajectory shifted dramatically on August 23, 1944, when King Michael I orchestrated a palace coup that arrested Antonescu and declared war on Germany. This switch to the Allied side altered Romania's post-war position. Soviet troops, having crossed the Prut River in April 1944 during their spring offensive, now entered as ostensible liberators rather than hostile conquerors. The August coup earned Romania recognition as a co-belligerent, allowing the nation to participate in post-war negotiations and avoid harsher occupation terms. Antonescu was executed in 1946 for war crimes, and the Communists seized power as they seized ministries, banned opposition, and abolished the monarchy in 1947 (Deletant, Hitler's Forgotten Ally).

Gheorghe Gheorghiu-Dej, the first communist leader of Romania, then imposed collectivization and nationalization, and in 1948 created the Securitate to direct arrests, surveillance, and show trials to maintain compliance amongst the populace (Deletant, *Communist Terror*). He also implemented the Soviet model of agrarian and industrial transformation which forced peasants to give up their land to collective farms under the state, which deprived many Romanians of their land, livelihood, and identity. State authorities also seized private factories, banks, and transport and transferred them to the Romanian Communist Party, also known as the Workers' Party.

After Gheorghiu-Dej died in 1965, powerful members of the Romanian Communist Party elected Nicolae Ceaușescu, a former shoemaker and member of the communist party, as General Secretary (Tismăneanu). Ceaușescu's early rule appeared comparatively liberal; he condemned the Warsaw Pact invasion of Czechoslovakia in 1968 (Petrescu and Petrescu 337) and obtained Western credit to fund projects (Jackson). Between 1968 and 1971 his government drew large Western loans for homes and factories as the regime's debt rose (Jackson). In July 1971, Ceaușescu delivered a speech known as the "July Theses" which reversed his comparatively liberal policies and imposed a harsh, authoritarian program (Siani-Davies). Ceaușescu replaced

independent intellectuals with party loyalists and he strengthened control over radio broadcasts and television, which only broadcast party-approved propaganda. After visits to China and North Korea in 1971, Ceaușescu's Stalinist inclinations were catalyzed and he considered importing the methods of indoctrination used during Mao's Cultural Revolution to Romania (Tismăneanu 206). Throughout the 1970s, Ceaușescu continued borrowing from Western creditors in an attempt to industrialize rapidly, a move which deepened Romania's national debt (Jackson). Export markets failed to fully absorb the new output and citizens lost access to basic goods, services, and resources such as electricity and food. Because of the national debt, Ceaușescu vowed to pay this off by 1990 and initiated many reforms. To repay creditors, Ceaușescu diverted much of Romania's production abroad and rationed food, fuel and electricity at near unlivable standards. Ceaușescu also launched a rural "systematization" program, a state planning scheme that sought to consolidate or eliminate villages and concentrate residents in standardized settlements (Bachman 79). By the late 1970s, Ceaușescu's regime relied on the Securitate for pervasive control and repression.

The Securitate Was Ceaușescu's Central Tool for Repressing Dissent and Fracturing Social Trust

Ceaușescu and his regime consolidated power and used the Securitate to oppress Romanian citizens politically. The Securitate's surveillance infiltrated every layer of Romanian society, from schools to telephones and television sets. The system fostered fear, isolated citizens from external influences, and twisted laws to justify violence and show trials. Each tactic only served to further shred communal bonds, stripping away a large component of what comprised Romanian identity.

To achieve his aims, Ceaușescu relied on the Securitate to monitor private communications and place microphones in homes (Bachman 301–02). The political police employed some 14,000 full-time agents and between 400,000 and 700,000 part-time informers (Stan). This created the sense that any home could become an audio surveillance post; consequently, private thoughts or beliefs could no longer be assumed safe among the populace. As such, this policy systematically eroded trust among Romanian citizens and severed the communal bonds that had historically provided resistance to authoritarian control by changing Romanian homes from sanctuaries of cultural transmission into performative spaces where genuine expression became impossible. Families could not pass down cultural stories, folk wisdom, or critical perspectives that had historically bonded generations out of fear of deviating from the regime's chosen narrative. Traditional Romanian hospitality of welcoming strangers and sharing stories suddenly became dangerous. This technological intrusion destroyed the open social fabric that had defined Romanian village life for centuries, where homes served as spaces for preserving oral traditions, folk songs, and communal bonds. When every conversation became potential evidence against the speaker, the Romanian language itself became impoverished. Metaphors, folk sayings, and cultural expressions that might be interpreted as criticism disappeared from daily speech, eliminating centuries of linguistic tradition that had

carried cultural meaning and collective memory. Children learned to mistrust even their parents' private thoughts, further severing the intergenerational transmission of Romanian cultural values. Beyond just surveillance, the Securitate wielded intimidation through arrests and other forms of abuse.

In the worst cases, Ceaușescu's regime and the Securitate utilized a multitude of intimidation tactics to create fear and oppress the population. Amnesty International's 1987 report documented the imprisonment, torture, house arrest, and harassment of government critics (Amnesty International). These abuses not only affected those directly targeted, but they also warned Romanians who considered defiance. The targeting of critics also resulted in widespread communal paranoia amongst the Romanian population and showed how intimidation tactics silenced independent Romanian ideas. Even minor issues carried serious risk and forced citizens to endure pain in silence. Rather than stage open protests, Romanians kept their opinions to themselves, a pattern that demonstrates Ceaușescu's oppression and intimidation which overall undermined civic unity among the populace. The randomness of arrests meant Romanians did not know how to act in a safe manner. This uncertainty forced Romanians to abandon not only political expression but the ordinary expression which made them Romanian. Fear based compliance replaced traditional Romanian expression and their concept of cinste, honor and integrity. Citizens became performers for the state instead of living true to their authentic selves. The Securitate's prison apparatus was also reported to administer lethal doses of toxic substances to political prisoners (Bachman 302). Such clandestine violence, oppression, and intimidation weakened Romanian communal trust and deepened the resentment that ultimately led to the upheaval of the regime.

When the Romanians refused to comply, Ceaușescu's regime used force to impose obedience. Livezeanu's transcript from 1989 of Ceaușescu's trial shows prosecutors accusing him of ordering the crackdown on the first Timișoara protest, an accusation he refused to answer ("Transcript of the Closed 'Trial'"). Tanks and gunfire forcefully and violently crushed protests across Romania and sent a powerful message. Ceaușescu's regime crushed every dialogue or protest for freedom and genuine Romanian identity, even at the cost of lives. The Association for Diplomatic Studies and Training also further confirmed that fear, and not loyalty, sustained public order in Ceaușescu's regime (Becker). This evidence shows that the Romanian population complied with the regime for survival rather than support. This resulted in a conditional stability amongst Romanian society, not a stability rooted in societal and cultural norms, but rather a fragile stability rooted in fear which eroded Romanians' national identity for the sake of their survival. The regime further suppressed the Romanian population by systematically isolating them from external influences.

Ceaușescu's regime and the Securitate restricted foreign contact and isolated Romania from the outside world. The Letter of the Six, a March 1989 open letter sent to Ceaușescu by six veteran Romanian Communist Party officials and later leaked to the international press, reported that the regime forbade citizen contact with the outside world (Apostol et al.). Ceaușescu's restrictions denied Romanians a point of comparison between their hardships and life in freer

societies abroad, limiting intellectual and cultural exchange. The ban on Western materials also restricted exposure to emigrant communities that could provide insight into pre-communist life. These restrictions created a cultural freeze in which Romanian identity had little room to evolve independently of the state. Contemporary analyses of Radio Free Europe broadcasting likewise describe the regime's effort to physically and intellectually isolate Romanians from the outside world (Johnson and Parta). By jamming foreign broadcasts and restricting contact, the regime narrowed the information available to the population and gave state narratives greater control over collective memory. The result was not only political isolation but cultural isolation: Romanians lost access to exchanges that had historically connected their language, literature, and arts to communities beyond the country's borders.

Ceaușescu's regime further suppressed Romanians through a systematic campaign of intimidation. The Letter of the Six reported that the regime cut telephones at will, breaching Article 34 of the constitution, which protected citizens' privacy (Apostol et al.). The Romanian population lost a basic line of personal and commercial communication. The falsely engineered dependency on the regime highlighted how Ceaușescu's regime utilized infrastructure as a weapon against Romanians to oppress them and minimize their cultural cohesion to prevent revolts. Telephone monitoring thus destroyed extended familial networks which stretched across borders. Romanian culture was partly maintained through its diaspora connections, but those ties were severed. The letter also detailed how authorities violated correspondence and cut off private communication (Apostol et al.). The regime absorbed private correspondence amongst the populace and invasive censorship blocked honest communication, and this surveillance highlights Ceaușescu's opposition and its erosion of Romanian civic identity. Under Gheorghiu-Dej, the regime had also utilized forced labor to construct the Danube–Black Sea canal (Deletant, Communist Terror). The forced labor symbolized the widespread systemic suppression of the communist regime as it dismantled the Romanian population's connections between each other. The Letter of the Six mentioned how its signatories risked their "liberty and even our lives" by speaking out against the regime (Apostol et al.). Romanians sought truth beyond official lines, and their hunger for knowledge revealed how deeply isolation had cut into Romanian life.

Ceaușescu's regime also politically oppressed the population under the Securitate's control with many crackdowns and trials which also seemed hypocritical at times. Ceaușescu's regime and the Securitate first denied the courts. Transcripts of Ceaușescu's trial from 1989 record how Ceaușescu dismissed judicial independence, as he stated "I only recognize the Grand National Assembly" ("Transcript of the Closed 'Trial'"). Ceaușescu's statement highlighted the lack of a real legal system in his regime. The Letter of the Six similarly argued that constitutional protections had effectively been suspended and that Romania no longer possessed a functioning legal system (Apostol et al.). Because Ceaușescu's regime offered little justice, the rule of law foundation necessary for communal trust had been absent as it corroded Romanian society both legally and culturally. No Romanian citizen on trial could expect a fair trial in any matter as trials resulted in whatever outcome the Securitate or Ceaușescu wanted. The false justice system

Ceaușescu's regime had signaled to Romanians that the regime existed to punish and control, not protect, and thereby it showed how false courts and political oppression eroded civic integrity. The collapse of legal protections weakened more than political rights: courts no longer functioned as shared institutions citizens could trust, reinforcing the fear and suspicion already produced by Securitate surveillance.

Ceaușescu's regime and the Securitate further oppressed the population through the crackdown on major demonstrations. In December 1989, troops and police opened fire on thousands of anti-government protesters in Timișoara (NATO). Romanian citizens who protested for basic rights were shot in Timișoara. The crackdown highlighted how the regime valued obedience over dialogue, which further corroded Romanians' national identity through the suppression of peaceful demonstrations. This episode shows how political violence took precedence over real governance, which highlighted how the regime's methods only eroded Romanians' social and cultural identity.

Ceaușescu's regime also utilized widespread violence against dissenters to maintain control and oppress the population which further eroded the populace's cultural identity. At the trial of Ceaușescu, prosecutors asked who "ordered the bloodbath in Timisoara," an accusation he refused to answer ("Transcript of the Closed 'Trial'"). The massacre sought to suppress protest through fear and tore Romanian society apart. Instead, the bloodshed intensified revolutionary zeal which highlighted not only the extent the regime went to oppress the population, but the collective anger the population had. The secret and brutal violence employed by the Securitate further fractured communal trust amongst the population. This secret violence only deepened the resolve to overthrow the regime.

Historian Dennis Deletant of University College London argued that the Securitate's most potent weapon was fear and that its success depended on fear inducing compliance (Deletant, "In and Out of Focus"). He argued that this fear kept citizens in line and helped the regime deter open resistance for years. Obedience from terror differs fundamentally from obedience which stems from loyalty and open consent. Ceaușescu silenced voices which hardened resentment and private resolve, so when fear fractured, a violent revolution erupted. The revolution demonstrated that repression hollowed out Romanian national identity and social bonds. Rather than unite them, the oppressive nature of Ceaușescu's regime undermined him, and in the end it did not sustain his rule.

Ceaușescu's Securitate shattered Romanian society and culture as surveillance silenced private Romanian thought while it punished dissent and cut Romanians off from the outside world. East Germany's Stasi police similarly eroded the community and culture of their own people through unofficial collaborators who reported on all areas of society and infiltrated opposition groups (Stasi Records Archive). Both surveillance services crushed trust amongst the population and simultaneously bred anger. In East Germany that anger culminated in 1989 when citizens tore down the Berlin Wall, while in Romania it culminated in the first gunshots in Timișoara. Both examples highlighted that coercion from authoritarian governments damaged society and eroded local cultures.

Ceaușescu's Debt-Driven Austerity Turned Modernization into Scarcity and Social Breakdown

Ceaușescu's fiscal policies created scarcity across Romania and turned agriculture into a debt-repayment engine while citizens faced deprivation at home. Three dynamics drove the crisis: export extraction to repay foreign debt, strict rationing and cuts to essential services, and industrial breakdown alongside visible privilege for the political elite. These policies did more than lower living standards. They redirected daily life toward obtaining food, heat, electricity, and wages, leaving less room for the family and community routines that had long helped sustain Romanian social life. The damage to identity came not from poverty alone, but from the way scarcity displaced the family, village, and workplace routines through which Romanian social life had been maintained.

Ceaușescu's economic failures stemmed from an unsustainable development strategy financed through heavy Western borrowing in the late 1960s and 1970s (Jackson). Factories expanded rapidly, but many lacked reliable supplies or markets for their output. The Letter of the Six warned in March 1989 that economic planning no longer functioned and that a growing number of factories lacked raw materials, energy, or markets (Apostol et al.). The letter also pointed to a striking imbalance in agriculture: private peasants working only 12 percent of the country's arable land produced 40 percent of its vegetables, 56 percent of its fruit, 60 percent of its milk, and 44 percent of its meat (Apostol et al.). Those figures exposed how poorly the centralized system was allocating Romanian resources. A modernization program intended to increase production instead left factories underused and households increasingly dependent on a state that could not reliably supply basic goods.

The burden of Ceaușescu's debt policy was especially visible in the food supply. The regime's demand for exports caused severe shortages of bread, meat, fruits, and vegetables (Bachman 59). Food produced inside Romania was diverted abroad to earn foreign currency while local markets faced recurring shortages. The Letter of the Six warned that continued food exports were threatening the biological existence of the nation (Apostol et al.). This policy also struck at a part of Romanian identity established earlier in the paper: village communities had historically tied land, family, and shared resources together. By removing locally produced food from domestic circulation and returning it through rationed, state-controlled channels, the regime weakened that relationship between land and community and made basic survival dependent on the same state that had created the shortages.

Ceaușescu's austerity imposed a severe burden on the population, as exemplified by chronic shortages of food, heat, and electricity. Many Romanians were forced to go without food, winter heat, and electricity for extended periods of time (Becker). By March 1989, the Letter of the Six stated that the government had failed for five consecutive winters to solve basic problems involving heating, lighting, transportation, and food (Apostol et al.). Long queues and recurring shortages became part of daily life (Bachman 59). The daily effort required to find food and keep homes warm redirected attention from ordinary family and community life toward survival. The time and energy required to obtain food and basic necessities reduced the space for family

gatherings, reciprocal support, and ordinary community routines that had historically reinforced social bonds.

Energy rationing made the scale of deprivation measurable. Authorities limited families to one forty-watt bulb per apartment and, during the severe winter of 1984-85, shut off heat and hot water (Bachman 59). The engineered deprivation turned homes into sites of anguish rather than sanctuaries. The same domestic spaces in which families had traditionally gathered, passed down stories, and maintained everyday customs became dominated by the immediate problem of keeping warm and finding light. These shortages therefore damaged more than material living standards; they weakened the household setting through which cultural habits and family memory were transmitted from one generation to the next.

Economic breakdown also weakened the relationship between workers and the state. A Romanian saying of the period captured the collapse in workplace trust: the regime pretended to pay workers, and workers pretended to work (Petrescu and Petrescu 330). In November 1987, wage reductions at the Steagul Roșu truck plant helped trigger the Brașov workers' revolt, which quickly expanded from an economic protest into open anti-government demonstrations (Petrescu and Petrescu 333). The episode showed how material deprivation could become political anger. When state employment no longer provided reliable wages or basic security, workplaces that were supposed to demonstrate socialist stability instead became centers of resentment against the regime.

The contrast between public austerity and elite privilege further undermined the regime's claim that sacrifice served a common national purpose. While ordinary citizens faced rationing, the political elite lived in palatial homes and had exclusive access to food and luxury goods unavailable to most citizens (Bachman 106). At the same time, Ceaușescu directed enormous resources toward monumental projects such as the Palace of the Parliament and the redevelopment surrounding it (Danta 175). The Letter of the Six charged that Bucharest's Civic Center, the country's largest multibillion-lei investment, had no public budget and that its cost had tripled because of Ceaușescu's repeated design changes (Apostol et al.). The disparity made austerity appear not as a shared national burden but as a policy imposed on the population while the leadership remained insulated from its effects. That visible inequality deepened resentment and weakened the sense of common obligation between citizens and the state.

Systematization and Propaganda Destroyed Physical Heritage and Recast Historical Memory as State Myth

Systematization was a state planning program that reorganized Romanian settlement space by consolidating rural populations and remaking urban areas through standardized socialist construction. Architectural historian Dana Vais notes that the term had entered Romanian urbanism earlier as a label for scientific planning before the socialist state politically instrumentalized it and applied it on a much larger scale (Vais 207). Under Ceaușescu, systematization became a mechanism for imposing a state-designed landscape over villages, neighborhoods, and religious sites. Because Romanian identity had long been rooted in those

local spaces, physical restructuring also became cultural restructuring: removing the places where communal life occurred weakened the settings in which local memory, religious practice, and family traditions were maintained.

Rural systematization posed this threat on a national scale. By 1988, between 7,000 and 8,000 villages were scheduled for eventual destruction or forced depopulation (Bachman 79). The Letter of the Six likewise warned that 7,000-8,000 villages were threatened and questioned why the state was trying to urbanize villages when it could not reliably provide heating, lighting, transportation, or food (Apostol et al.). The letter also described systematization as the forced removal of peasants into three-story apartment blocks, directly separating villagers from the households and village spaces around which local life had developed (Apostol et al.). Romanian villages were not merely collections of houses. As established earlier, they were places where extended families shared land, church parishes organized communal life, and stories and customs passed between generations. Treating thousands of villages as expendable therefore threatened the social structure that carried those practices forward.

Systematization also reshaped Bucharest during the 1980s. In the redevelopment of the capital's center, about 40,000 people were forced to leave their homes with only twenty-four hours' notice, while roughly eight square kilometers in the heart of Bucharest were destroyed (Bachman 79). Danta estimates that the demolished zone accounted for roughly 20-25 percent of historic Bucharest (Danta 175). Among the lost sites were Ypsilanti Palace, the eighteenth-century Brâncovenesc Hospital, and eighteen churches and monasteries (Lykiardopol 27). These were not interchangeable buildings: they were physical anchors to neighborhood history, religious practice, and shared memory. Bachman explicitly notes that the destruction of churches and villages severed cultural and historical links to the past and threatened community bonds and group autonomy (Bachman 80). By removing both residents and the spaces around which communities had formed, systematization disrupted the places through which Romanian identity was lived and remembered.

Physical demolition worked alongside propaganda. As older visual reminders disappeared, state media increasingly supplied an official story of Romanian history. By the mid-1980s, Ceaușescu's personality cult had reached enormous proportions, and official propaganda placed him in the lineage of Romanian historical rulers (Bachman 59). Danta's analysis of Bucharest shows how the destruction of visual reminders of the past cleared space for a manufactured historical narrative (Danta 175). The regime therefore did not simply remove the past from the landscape; it also tried to determine which past remained visible and how that past should be interpreted. This combination of demolition and propaganda gave the state greater power over collective memory itself.

One of the clearest examples of this manipulation was the regime's use of medieval rulers. Michael the Brave (Mihai Viteazul), the Wallachian prince whose campaigns in 1600 briefly brought Wallachia, Moldavia, and Transylvania under one ruler and who later became a symbol of Romanian unification, was promoted through schoolbooks and films as part of the national-communist narrative (Ghender 178-80). Vlad Țepeș, the fifteenth-century Wallachian

prince remembered for resisting Ottoman power and enforcing centralized authority, was also reinterpreted by communist historians as a defender of independence and strong rule, inviting parallels with Ceaușescu's leadership (Light 14). Ceaușescu was presented alongside exceptional historical leaders rather than as a break from them. Schoolbooks, films, posters, and official historical writing thus did more than praise Ceaușescu: they placed him inside familiar national stories and redirected inherited symbols toward loyalty to the regime. Romanian historical memory remained present, but the state increasingly controlled the meaning attached to it. State control over media and cultural production reinforced that historical narrative. Throughout the 1970s and 1980s, the media were reorganized under increasingly stringent party control, while writers and other intellectuals faced political control over their careers and expression (Bachman 104). Films, television, publishing, and education therefore operated within boundaries set by the state. When alternative interpretations of Romanian history and culture could not circulate freely, public culture narrowed to what the regime was willing to approve. The damage to identity came from that narrowing: Romanian traditions and historical figures were not erased completely, but their public meaning was filtered through the political needs of Ceaușescu's government.

Religious policy placed similar pressure on another central source of Romanian identity. Although the state formally recognized religious groups, it subjected them to strict controls, prohibited religious activity outside designated buildings, and restricted religious education and materials (Bachman 120). Systematization compounded those restrictions by physically removing churches and attempting to replace their communal role. Bachman notes that new Houses of Culture were intended to replace churches as focal points of community life and that, by 1989, many churches had already been destroyed (Bachman 80). Replacing parish churches with state civic institutions changed more than the skyline. The policy weakened spaces where worship, family rites, local memory, and community gathering had historically overlapped, shifting communal life away from inherited religious institutions and toward institutions controlled by the state.

Conclusion

Ceaușescu's policies damaged the institutions and spaces through which Romanian identity was lived. Political repression made private speech dangerous and weakened interpersonal trust. Debt-driven austerity redirected daily life from ordinary family and community routines toward the search for food, heat, electricity, and wages. Systematization removed villages, churches, neighborhoods, and historic buildings that anchored communal memory, while propaganda recast national history and familiar rulers around Ceaușescu's personality cult. Together, these policies did not simply make the regime unpopular; they weakened social trust, local autonomy, cultural memory, and the everyday institutions that connected Romanians to one another. Ceaușescu's forced vision of a "modern" Romania helped produce obedience for a time, but not durable loyalty. When fear fractured in 1989, the regime had little legitimacy left to hold it together.

Works Cited

Amnesty International. Romania: Human Rights Violations in the Eighties. Amnesty International, 1987. EUR 39/02/87, www.amnesty.org/en/documents/eur39/002/1987/en/. Accessed 20 Aug. 2026.

Apostol, Gheorghe, et al. "Letter of the Six, March 1989." World History Commons, worldhistorycommons.org/letter-six-march-1989. Accessed 20 Aug. 2026.

Bachman, Ronald D., editor. Romania: A Country Study. Federal Research Division, Library of Congress, 1991, www.loc.gov/item/90006449/. Accessed 20 Aug. 2026.

Becker, Frederick. "The 1989 Romanian Revolution and the Fall of Ceausescu." Interview. Association for Diplomatic Studies and Training, Oct. 2015, adst.org/2015/10/the-1989-romanian-revolution-and-the-fall-of-ceausescu/. Accessed 20 Aug. 2026.

Danta, Darrick. "Ceaușescu's Bucharest." Geographical Review, vol. 83, no. 2, 1993, pp. 170–82. JSTOR, https://doi.org/10.2307/215255.

Deletant, Dennis. Communist Terror in Romania: Gheorghiu-Dej and the Police State, 1948–1965. Hurst, 1999.

Deletant, Dennis. Hitler's Forgotten Ally: Ion Antonescu and His Regime, Romania 1940–1944. Palgrave Macmillan, 2006, https://doi.org/10.1057/9780230502093.

Deletant, Dennis. "In and Out of Focus: An Academic in the Sights of the Securitate 1965 to 1989." UCL European Institute, University College London, 2 Dec. 2015, www.ucl.ac.uk/european-institute/analysis/2015-16/dennis-deletant. Accessed 20 Aug. 2026.

Ghender, Flavius. "Michael the Brave: The Construction of National Hero Myth in Cinematography." Journal of Humanistic and Social Studies, vol. 8, no. 1, 2017, pp. 177–85, jhss.ro/downloads/15/articles/12%20Ghender.pdf. Accessed 20 Aug. 2026.

International Commission on the Holocaust in Romania. Final Report. Chaired by Elie Wiesel, Polirom, 2004.

Jackson, Marvin R. "Romania's Debt Crisis: Its Causes and Consequences." East European Economies: Slow Growth in the 1980s, vol. 3, Joint Economic Committee, U.S. Congress, U.S. Government Printing Office, 1986, pp. 489–542.

Johnson, A. Ross, and R. Eugene Parta, editors. Cold War Broadcasting: Impact on the Soviet Union and Eastern Europe. Central European UP, 2010.

Kligman, Gail. The Wedding of the Dead: Ritual, Poetics, and Popular Culture in Transylvania. U of California P, 1988.

Light, Duncan. "The Status of Vlad Țepeș in Communist Romania: A Reassessment." Journal of Dracula Studies, vol. 9, no. 1, 2007, pp. 9–23, https://doi.org/10.70013/76h1i2j3.

Lykiardopol, Matei. "The Mutilation of Bucharest." The UNESCO Courier, Jan. 1991, pp. 26–29.

North Atlantic Treaty Organization. "Troops and Police Open Fire on Thousands of Anti-Government Protesters in the Romanian Town of Timisoara." NATO, 20 Dec. 1989,

www.nato.int/en/news-and-events/articles/news/1989/12/20/troops-and-police-open-fire-on-thousands-of-anti-government-protesters-in-the-romanian-town-of-timisoara. Accessed 20 Aug. 2026.

Petrescu, Cristina, and Dragoș Petrescu. "Resistance and Dissent under Communism: The Case of Romania." Totalitarismus und Demokratie, vol. 4, no. 2, 2007, pp. 323–46.

Siani-Davies, Peter. The Romanian Revolution of December 1989. Cornell UP, 2005.

Stahl, Henri H. Traditional Romanian Village Communities: The Transition from the Communal to the Capitalist Mode of Production in the Danube Region. Translated by Daniel Chirot and Holley Coulter Chirot, Cambridge UP, 1980.

Stan, Lavinia. "Inside the Securitate Archives." Wilson Center, 4 Mar. 2005, www.wilsoncenter.org/article/inside-the-securitate-archives. Accessed 20 Aug. 2026.

Stasi Records Archive. "The Unofficial Collaborators (IM) of the MfS." Federal Archives, www.bundesarchiv.de/en/stasi-records-archive/education/what-was-the-state-security/the-unofficial-collaborators-of-the-mfs/. Accessed 20 Aug. 2026.

Tismăneanu, Vladimir. Stalinism for All Seasons: A Political History of Romanian Communism. U of California P, 2003.

"Transcript of the Closed 'Trial' of Nicolae and Elena Ceausescu, December, 1989." Presented by Irina Livezeanu. Making the History of 1989, item 690, Roy Rosenzweig Center for History and New Media, chnm.gmu.edu/1989/items/show/690. Accessed 20 Aug. 2026.

Vais, Dana. "Systematization: A Key Term in 20th-Century Romanian Urbanism." Urban Planning, vol. 7, no. 1, 2022, pp. 207–22, https://doi.org/10.17645/up.v7i1.4791.

Conscience and Constraint: Lillian Hellman's Testimony and the Limits of Cold War Resistance By Evelyn Cai

Abstract

This paper examines the 1952 testimony of playwright Lillian Hellman before the House Un-American Activities Committee (HUAC) as a case study in the limits of individual resistance under Cold War political pressure. Rather than treating Hellman's refusal to name others as a simple act of courage, this paper argues that her testimony was a strategically constructed yet structurally constrained response to a coercive system that operated outside the safeguards of the criminal justice system. The analysis proceeds in three stages. First, it traces how HUAC's institutional power depended on covert collaboration with the FBI and on a parallel system of public exposure that bypassed the protections of a criminal trial. Second, it situated Hellman within Hollywood's culture of suspicion, in which studio executives and anti-communist politicians used accusations of subversion to discipline labor organizing and cultural dissent. Third, it reconstructs Hellman's own strategy of qualified cooperation, comparing her approach to that of contemporaries who chose different paths before the committee. Drawing on committee records, biographical scholarship, and contemporary press coverage, the paper concludes that Hellman's testimony reveals a paradox at the heart of Cold War resistance: the same legal and rhetorical strategies that allowed her to avoid contempt of Congress also exposed her to sustained professional and financial punishment, demonstrating that resistance within an authoritarian climate was less a matter of individual courage than of strategy shaped and limited by the very system it sought to oppose.

Introduction

"I will not cut my conscience to fit this year's fashion," playwright and screenwriter Lillian Hellman declared in her 1952 letter to the House Un-American Activities Committee (HUAC), as Cold War tensions in the United States escalated to the point where dissent was treated as a threat to national security (Haynes 411). Hellman, one of the most prominent American playwrights of the twentieth century, was best known for works such as The Children's Hour (1934) and The Little Foxes (1939), which had established her as a major cultural figure whose political views were widely scrutinized (Wright 83). In this climate, political suspicion became embedded in federal policy and cultural life, with institutions like HUAC—a congressional committee established in 1938 to investigate alleged subversive influence in American life—transforming testimony into a tool of coercion (Kessler-Harris 112-113; Morgan 114). Cold War political pressures both compelled Hellman's appearance before HUAC and shaped her refusal to cooperate, turning her testimony into a strategically constructed yet structurally constrained act of resistance with lasting consequences whose significance remains debated, since she avoided legal punishment but suffered enduring professional and financial damage. These pressures first emerged through the institutional power of HUAC and its collaboration with federal agencies, then intensified within Hollywood's culture of suspicion and blacklisting, before shaping Hellman's personal strategy and the outcome of her testimony.

The Institutional Power of HUAC

HUAC's power over witnesses emerged from a coordinated system that combined federal policy and covert intelligence, shaping both the process and outcome of its investigations. Unlike a court of law, HUAC held no formal judicial authority, yet it wielded significant coercive power: it could subpoena private citizens, hold them in contempt of Congress for non-cooperation, and expose their alleged affiliations publicly, effectively destroying careers without a criminal conviction (Haynes 408-409). Before Senator Joseph McCarthy's rise, policies such as Harry Truman's loyalty program and the Attorney General's list had already institutionalized suspicion and enabled widespread investigation of alleged subversives, blurring the line between political dissent and disloyalty (Wertheim 211-212). HUAC expanded within this framework, building on earlier investigations that normalized public fear and legitimized scrutiny of cultural figures (Morgan 531). Its pursuit of the Hollywood Ten, for instance, turned non-cooperation into a public spectacle, equating refusal with guilt and reinforcing the committee's authority (Morgan 531-532). Simultaneously, the Federal Bureau of Investigation (FBI) directly supported HUAC's work by supplying confidential files, names, and background data on suspected affiliations, leading HUAC to "document that the recalcitrant witnesses were Communists only because of FBI assistance," exposing the committee's dependence on surveillance (Theoharis 156-157, 164). Practices that allowed HUAC to present evidence without disclosing the FBI as its source obscured their collaboration, masking the extent of the executive branch's involvement in congressional investigations (Theoharis 160-163). Rather than seeking neutrality, HUAC entered hearings having already presumed guilt, using FBI intelligence to confirm suspicions it had formed in advance (Theoharis 156-159). Thus, HUAC's authority rested on a broader network that preconditioned both the terms of testimony and the consequences of resistance.

HUAC's coercive power is best understood not as a substitute for the criminal justice system but as a parallel one, operating by different rules and evading its safeguards. A witness before the committee enjoyed none of the protections that would accompany a criminal trial—no presumption of innocence, no requirement that guilt be proven beyond a reasonable doubt, and no meaningful opportunity to cross-examine accusers—yet faced consequences, in lost employment and public disgrace, that could be more severe and more immediate than many criminal penalties. This asymmetry is what made the committee's staged, televised, and press-covered hearings so effective as instruments of political discipline: conviction in the court of public opinion required none of the evidentiary standards that would have been necessary to convict a witness of an actual crime. HUAC's investigators understood this distinction and exploited it deliberately, treating publicity itself as the punishment rather than as merely incidental to a formal proceeding. For a witness like Hellman, this meant that even a technically successful legal outcome—avoiding contempt, avoiding prison—could not undo the reputational and economic injury inflicted by the hearing's public character. The committee's power, in other words, rested less on the sanctions it could formally impose than on the informal machinery of

blacklisting and stigma that its hearings triggered, a machinery entirely outside its own legal jurisdiction but wholly dependent on the legitimacy its congressional status supplied.

Hollywood's Culture of Suspicion and the Blacklist

By the late 1940s, the pressures shaping Hellman's testimony had developed within an industry where anti-communist politicians and profit-minded studio executives found common cause—the former using Hollywood's visibility to dramatize the Communist threat, the latter using accusations of subversion to suppress labor organizing—transforming political suspicion into an inseparable part of professional life. The Screen Writers Guild (SWG) stood at the center of these tensions, as studio executives increasingly accused the guild of Communist influence to justify limiting its power (Kessler-Harris 112-113; Morgan 517). This framing transformed labor disputes into ideological battles, allowing studios to present economic demands as political threats and to recast union activity as subversive (Kessler-Harris 104). HUAC reinforced this arrangement by granting studios political cover as cooperating with the committee shielded executives from scrutiny and gave them a mechanism to discipline politically active employees (Kessler-Harris 254). Communist-affiliated organizations such as the American Writers Congress received public criticism from prominent figures such as novelist Thomas Mann, who, after addressing the Congress in 1939, remarked that it concerned itself "too much about politics and not enough about literature." (Morgan 171) The tension between visibility and suspicion made later accusations more potent, as individuals who had once participated openly were reinterpreted as a part of a hidden ideological threat. In addition, the secretive structure of communist party organization in Hollywood allowed investigators to portray Communism as covert and conspiratorial (Eckstein 429-430). Thus, testimony was a self-reinforcing mechanism, as naming others supplied HUAC with expanding lists of suspects, producing what Hellman later described in a 1961 Harvard lecture as "confessions of sin that never happened." (Kessler-Harris 257) Within this environment, blacklisting operated both formally and informally: studios denied employment based on suspicion alone, and workers who refused to cooperate with HUAC found themselves unable to find work across the industry regardless of whether formal charges were ever brought against them (Eckstein 424-425; Wertheim 212). This machinery had a concrete, semi-official infrastructure behind it: the right-wing newsletter Counterattack compiled and circulated lists of suspected Communists working in film and theater, and the FBI supplied confirming information that studios then used to justify quietly excluding named individuals from future work (Kessler-Harris 253).

The dual public-private character of this system was what made it so difficult to resist. Had blacklisting depended solely on government action, it might have been challenged in court as a violation of due process; had it depended solely on private industry discretion, it might have lacked the aura of national-security urgency that made resistance seem unpatriotic. By combining a congressional investigation with privately compiled clearance lists and voluntary studio compliance, the system achieved a form of punishment that operated almost entirely outside constitutional protections while still drawing its authority from the state. No court could order a

studio to rehire a blacklisted writer, because no studio was legally obligated to hire anyone in the first place, and no court could compel HUAC to prove its suspicions, because the committee never claimed to be adjudicating guilt. The result was a mechanism nearly immune to legal challenge, which is precisely why avoiding contempt charges, as Hellman managed to do, offered so little practical protection against its effects.

Hellman's Testimony and Strategic Resistance

A celebrated playwright whose political convictions were well documented, Hellman's experiences in leftist politics made her a vulnerable target for scrutiny in this tense climate of heightened anti-communist investigation. Her political views were rooted in a strong sympathy for the working class and a deep distrust of the wealthy, leading her to support ideals of class equality and racial justice—convictions she had held and expressed publicly since at least the 1930s. Her outlook reflected what critics later described as her "naive politics," a belief that a more just and equitable society was worth defending even within flawed political systems (Kessler-Harris 102). This perspective later informed her critique of anti-Communist intellectuals, whom she viewed as complicit in prioritizing ideological loyalty over the defense of democratic rights (Falk and Hellman 97-98). Such convictions had led her, by the late 1930s, to engage publicly with leftist causes, including public support for Stalin's show trials in April 1938, which placed her within politically active cultural circles (Wright 215-216; Morgan 171; Kessler-Harris 102). As anti-communist investigations intensified, Hellman responded publicly with anger, writing in the SWG publication Screen Writer in December 1947 that the hearings were "sickening, immoral and degraded" and condemning cooperating witnesses as "craven men." (Kessler-Harris 251) These statements, made on the record, ensured that when investigations expanded into Hollywood, Hellman would be among the first targets. Meanwhile, testimony from Martin Berkeley in September 1951, who claimed Hellman attended a meeting organizing the Communist Party in Hollywood at his house in June 1937, reinforced suspicions and placed her within an increasingly inspected network (Wright 129). Hellman's position was also shaped by the fact that she was one of relatively few women to appear before the committee as a prominent, independent political actor rather than as the wife or associate of a male target. Press coverage of her case often folded assessments of her politics together with commentary on her appearance and manner—her composure, her clothing, her voice—in a way rarely applied to the men of the Hollywood Ten. This scrutiny cut two ways: it exposed her to a kind of judgment her male counterparts largely escaped, but it also gave her a tool the men lacked, since a carefully managed public image could soften the threat a defiant woman was assumed to pose. Together, her past affiliations, public statements, and external accusations positioned Hellman as a figure defined by suspicion, whom the committee had already formed conclusions about before her testimony.

When Hellman received her subpoena on February 21, 1952, summoned to testify about her alleged Communist affiliations, she faced a narrow and constrained set of options. Working with her lawyer Joseph Rauh, she developed a strategy of limited cooperation: she would answer

questions about herself and her own activities, but refuse to identify others, invoking the Fifth Amendment selectively to avoid contempt charges (Falk and Hellman 99). This middle position was legally precarious as the Fifth Amendment's protections could be lost if a witness answered some questions and not others on the same subject, leaving her vulnerable if the committee pressed further (Haynes 409-410; Wright 248). This vulnerability was not merely theoretical. A year earlier, in Rogers v. United States, the Supreme Court had upheld a contempt citation against a grand jury witness who, having voluntarily answered questions about her role in the Communist Party of Denver, was held to have thereby waived her Fifth Amendment protection with respect to related questions, including the identity of another party member (Rogers v. United States 367). The decision effectively held that a witness could not selectively invoke the privilege once she had begun to answer on a given subject—precisely the strategy Rauh was constructing for Hellman. Hellman and Rauh were therefore not simply choosing a cautious middle path; they were gambling that HUAC, unlike a grand jury bound by a recent and directly applicable precedent, would treat her partial cooperation as sufficient rather than press the waiver argument to its logical conclusion. Rauh had initially hoped to strengthen her defense by pointing to moments when she had publicly disagreed with Communist positions, including criticism from New Masses, an American Marxist magazine that had attacked her play Watch on the Rhine. Hellman rejected this approach, unwilling to attack the Communist Party at a moment when it faced persecution (Wright 245). An earlier draft of her statement, prepared under Rauh's direction in late April, had gone further than the letter she ultimately sent: it conceded that she had been "wrong about the Communist party" and had joined out of misplaced idealism, language Hellman evidently judged too concessive once the final version was drafted a few weeks later (Kessler-Harris 260). This decision preserved her moral integrity but narrowed her legal options further, since it removed the strongest evidence that she had maintained independent judgment rather than blind party loyalty. She was also determined to avoid both imprisonment and the act of naming others, while protecting her public image—three goals that pulled against each other, since the safest legal path, full cooperation, would have required betraying the very principle she most wanted to defend (Wright 249). As a result, Hellman carried into her testimony a strategy that was morally coherent but legally fragile, depending on protections the committee was not obliged to honor.

Hellman's testimony on May 21, 1952 translated her carefully constructed strategy into a controlled public performance, in which she managed both her answers and her image simultaneously. Before taking her seat, she had sent a letter to the committee on May 19, framing her position as a matter of personal ethics rather than political allegiance, condemning the act of naming others as "inhuman and indecent and dishonorable" while invoking "simple rules of human decency" and a "good American tradition." (Wright 247) The letter also insisted on her own clear conscience, stating that "there is nothing in my life of which I am ashamed," and drew a deliberate distinction between her willingness to answer for herself and her refusal to answer for others: though she acknowledged holding a constitutional privilege to decline any question about her political opinions, activities, and associations, she wrote that she did "not wish to claim

this privilege" on her own behalf, reserving her refusal for questions that would force her to inform on others (Wright 246-247). By grounding her refusal in moral rather than political language, she shifted the terms of the hearing away from Communism and toward a universally recognizable standard of honor—a reframing she anticipated would matter as much to the press and public as to the committee itself (Falk and Hellman 99). During the hearing, she answered questions about her own activities and past affiliations while invoking the Fifth Amendment when asked to identify others (Haynes 409-410; Wright 251). For instance, when committee counsel Tavenner asked whether she had ever been a member of the Communist Party, Hellman refused to answer on "the same grounds," invoking the Fifth, yet moments earlier she had answered "No, sir" when asked if she was currently a member (Wright 251). This pattern effectively acquiesced to the committee's right to question her while still protecting others from implication (Wright 253). Her letter was entered into the record, and Rauh distributed copies to the audience, ensuring her moral framing accompanied the proceedings in real time (Kessler-Harris 261-262). Hellman also purchased a Balmain dress, hat, and gloves for the occasion—a deliberate choice that signaled her awareness of the broader audience and her intention to stage the hearing as a public event, not merely a legal one (Wright 250). As a playwright, Hellman was uniquely equipped to recognize that a congressional hearing was, among other things, a piece of stagecraft, complete with a script the committee expected witnesses to follow: contrition, cooperation, and the naming of names. Her refusal to perform that script, replaced with a competing performance of her own devising—the composed woman in a couture gown, reading a carefully drafted statement of conscience rather than answering spontaneously—was itself a form of dramaturgy, one that borrowed the tools of her professional craft and turned them against the forum meant to discredit her. The hearing's brevity only heightened this effect: rather than allowing the committee to draw out a lengthy, grinding interrogation of the kind endured by the Hollywood Ten, Hellman's letter closed off most avenues of extended questioning before they could open, compressing the encounter into something closer to a single scene than a trial. The hearing lasted one hour and seven minutes (Wright 251). Unlike the Hollywood Ten, who had refused to cooperate at all and been cited for contempt, Hellman's selective compliance kept her out of jail—though it did not protect her from the professional consequences that followed.

Reception, Consequences, and Legacy

Hellman's testimony generated a public narrative shaped as much by press framing and her own strategic self-presentation as by what she had actually said before the committee. Her letter to HUAC had functioned as a calculated public document from the start, crafted with the expectation that it would circulate and influence how audiences interpreted her stance (Haynes 409-410). The press responded favorably, with The New York Times reporting on May 22 that she had refused to "hurt innocent people...to save \[herself\]," foregrounding her ethics over her political affiliations (Haynes 411; “Lillian Hellman Balks House Unit”). Journalist Murray Kempton similarly praised her actions as a "courageous act of conscience" and a "worthy...code

of honor" in the New York Post on May 26, framing her stance as aligned with American values (Kessler-Harris 262). However, not all reactions were so generous as novelist and critic Mary McCarthy later dismissed her writing entirely, declaring that "every word she writes is a lie, including and and the," suggesting that the favorable narrative Hellman constructed around her testimony was far from universally accepted (Kessler-Harris 323). The gap between Kempton's and McCarthy's assessments reflected a deeper fracture among American intellectuals over how to respond to Communism, one that predated and outlasted the hearing itself. Liberal anti-Communists such as McCarthy, many of them former or disillusioned leftists associated with journals like Partisan Review, viewed figures like Hellman as insufficiently repentant fellow travelers whose refusal to fully disavow the Communist Party—rather than merely refusing to name others—compromised any claim to principled resistance. Hellman's admirers, by contrast, argued that this distinction between refusing to inform and refusing to renounce was exactly the point: one could reject the coercive premise of the hearings without being required to prove one's own innocence to the committee's satisfaction. This disagreement did not end with Hellman's testimony. It resurfaced publicly nearly three decades later, in early 1980, when McCarthy repeated her charge that Hellman was fundamentally dishonest during a television interview with Dick Cavett, prompting Hellman to file a libel suit that remained unresolved until her death in 1984—a reminder that the argument over what her testimony had actually proven continued to shape her reputation for the rest of her life (“Lillian Hellman Sues Mary McCarthy for Libel”). Hellman's established public reputation nonetheless helped shield her from harsher consequences, as prosecuting her risked backlash and made the committee appear excessive (Falk and Hellman 99-100). Beyond journalism, revivals of her play The Children's Hour on Broadway were read by theater audiences as direct commentary on HUAC's culture of accusation and reputational destruction, extending her moral framing beyond the hearing itself (Wertheim 218). Together, press coverage and cultural reception shaped how Hellman's testimony was understood and ultimately softened the consequences she faced.

Although Hellman avoided the most severe legal consequences of her testimony, the aftermath reveals that Cold War political pressures imposed lasting professional, financial, and psychological costs that her strategy could not prevent. By pleading the Fifth selectively, she was neither jailed nor fully blacklisted, yet she still lost professional opportunities and faced significant economic decline (Falk and Hellman 100). Hellman herself was caught in the same apparatus described above: her name continued to circulate on Counterattack-style lists after 1952, and by the mid-1950s she found herself effectively excluded from Hollywood screenwriting work irrespective of the legal outcome of her testimony (Kessler-Harris 253). Her own account in Scoundrel Time captures this instability directly, recalling that she had "enough money for a few years and then \[she\] didn't have any" and that "the farm had to be sold," indicating that her subpoena set off a financial decline from which she did not quickly recover (Hellman 114). Stigma persisted regardless, as she continued to be labeled as a "known communist," her reputation remaining shaped by public suspicion (Kessler-Harris 265). Yet this narrative was not absolute. A few weeks after her appearance, Hellman performed publicly at the

YMHA in Manhattan and received a standing ovation from theater audiences, still wearing the same Balmain dress, demonstrating that significant cultural support survived her testimony (Wright 256). Private praise from theater critic Brooks Atkinson similarly affirmed her stance, with Atkinson writing that her "attitude has a basic moral that every lover of the American system must adhere to," reinforcing her image as both principled and patriotic (Kessler-Harris 262). That such support coexisted with financial ruin, professional stigma, and persistent fear reveals the psychological toll her experience exacted—one could be publicly celebrated and privately devastated simultaneously (Kessler-Harris 263). Thus, her resistance shaped her public image but could not prevent its lasting consequences.

The range of possible responses to a HUAC subpoena is thrown into relief by comparison with a contemporary whose choice diverged sharply from Hellman's: the director Elia Kazan, who appeared before the committee twice in 1952, first refusing to identify former associates in January and then, in April, naming eight one-time colleagues from the Group Theatre who had belonged to the Communist Party alongside him (“Elia Kazan”). Unlike Hellman, Kazan faced comparatively little professional risk in refusing a second time: his standing on Broadway, largely untouched by the Hollywood blacklist, would likely have survived non-cooperation, and his decision to name names was widely read afterward as a matter of choice rather than necessity. The contrast sharpens what made Hellman's position distinctive. She pursued a middle course precisely because full cooperation was available to her too, and because the witnesses who chose it, like Kazan, generally emerged with their careers intact, while those who refused entirely, like the Hollywood Ten, generally did not. Her strategy was calibrated to occupy the narrow space between these two outcomes, and the fact that so few witnesses managed to hold that space credibly is itself evidence of how difficult, and how exceptional, her performance before the committee was.

Conclusion

Ultimately, the Cold War political framework that empowered HUAC also defined the limits of resistance within it, and Lillian Hellman's testimony illustrates both dimensions. The passage of the Smith Act in 1940 and its affirmation by the Supreme Court in Dennis v. United States established a precedent that treated political belief as a potential threat to national security, reinforcing a system in which suspicion alone justified punishment (Kessler-Harris 114). Within this context, Hellman's testimony was not a singular act of defiance, but the product of pressures operating across institutions, industries, and individual lives. The same system that empowered HUAC also reshaped Hollywood into a space where professional survival depended on political conformity, leaving witnesses with only constrained choices. Whether her testimony should be judged a success depends entirely on the standard applied: measured against the fate of the Hollywood Ten, who were cited for contempt and imprisoned, her middle position was a calculated achievement—she avoided jail, avoided naming others, and shaped a public narrative that cast her as principled rather than evasive. Measured against her own life before the subpoena, however, the costs were real and lasting—a farm sold, a career diminished, and a

reputation that remained contested for decades. This pattern of policing culture through accusations of disloyalty was not confined to the United States: comparable campaigns in the Soviet bloc, such as Andrei Zhdanov's purge of the cultural intelligentsia in the late 1940s, pursued the same goal of marginalizing writers and artists judged ideologically suspect, suggesting that Hellman's ordeal belonged to a broader Cold War pattern of cultural coercion rather than an anomaly of American anti-communism alone (Johnston 304). Yet the structural differences between the two systems matter as much as their shared logic. Zhdanov's purges were executed directly by the state, through Communist Party organs with the power to strip artists of their union membership, their access to publication, and in some cases their liberty. HUAC could do none of these things on its own authority; its power depended entirely on voluntary compliance from private employers who found it more profitable to cooperate than to resist. This difference did not make the American system less effective—if anything, the illusion of private choice made it harder to challenge, since no single actor could be held responsible for the cumulative effect of a boycott assembled from thousands of individually defensible decisions—but it does help explain why Hellman's resistance, however constrained, remained possible at all. She could not have contested a factory dismissal notice in Zhdanov's Soviet Union through an open letter and expected it to circulate freely, let alone to shape her public reputation favorably; the same qualified freedoms that exposed her to sustained economic punishment also gave her the platform from which to contest it. Her experience demonstrates that under Cold War political pressure, resistance was not simply an act of courage, but a strategy shaped and limited by the system it sought to oppose.

Works Cited

Eckstein, Arthur. "The Hollywood Ten in History and Memory." Film History, vol. 16, no. 4, 2004, pp. 424-436. JSTOR, https://www.jstor.org/stable/3815610.

"Elia Kazan ~ About Elia Kazan." American Masters, PBS, accessed Aug. 2026, https://www.pbs.org/wnet/americanmasters/elia-kazan-about-elia-kazan/642/.

Falk, Richard A., and Lillian Hellman. "Comment: 'Scoundrel Time': Mobilizing the American Intelligentsia for the Cold War." Performing Arts Journal, vol. 1, no. 3, 1977, pp. 97-102. JSTOR, https://doi.org/10.2307/3245252.

Haynes, John Earl. "Hellman and the Hollywood Inquisition: The Triumph of Spin-Control over Candour." Film History, vol. 10, no. 3, 1998, pp. 408-414. JSTOR, https://www.jstor.org/stable/3815233.

Hellman, Lillian. Scoundrel Time. Little, Brown, 2000.

Johnston, Gordon. "Revisiting the Cultural Cold War." Social History, vol. 35, no. 3, 2010, pp. 290-307. JSTOR, https://www.jstor.org/stable/27866662.

Kessler-Harris, Alice. A Difficult Woman: The Challenging Life and Times of Lillian Hellman. Bloomsbury Press, 2012.

"Lillian Hellman Balks House Unit." The New York Times, 22 May 1952, p. 15, https://timesmachine.nytimes.com/timesmachine/1952/05/22/issue.html.

"Lillian Hellman Sues Mary McCarthy for Libel." HISTORY, A&E Television Networks, last modified 14 Feb. 2020, https://www.history.com/this-day-in-history/february-14/lillian-hellman-sues-mary-mccarthy.

Morgan, Ted. Reds: McCarthyism in Twentieth-Century America. Random House, 2003.

Rogers v. United States. 340 U.S. 367, Supreme Court of the United States, 1951.

Theoharis, Athan G. Chasing Spies: How the FBI Failed in Counterintelligence but Promoted the Politics of McCarthyism in the Cold War Years. Ivan R. Dee, 2002.

Wertheim, Albert. "The McCarthy Era and the American Theatre." Theatre Journal, vol. 34, no. 2, 1982, pp. 211-222. JSTOR, https://doi.org/10.2307/3207451.

Wright, William. Lillian Hellman: The Image, the Woman. Simon and Schuster, 1986.

Global Energy and International Relations: A Study of Energy Security, Geopolitics and the Global Energy Transition By Toshani Padarthi

Abstract

Energy has shaped international relations since the Industrial Revolution, influencing how states build their economies, exercise power and pursue their foreign policy objectives. Although the global shift towards low-carbon energy is often presented as a move away from fossil-fuel geopolitics, this paper argues that it is changing the nature of geopolitical competition rather than replacing it. Oil and natural gas remain central to economic stability and national security, while competition over critical minerals, manufacturing and clean-energy technologies is creating new strategic dependencies.

This research examines the relationship between global energy and international relations through a qualitative approach based on a structured literature review, policy analysis and secondary data from academic research, international organisations, government publications and industry reports. It traces the historical development of energy geopolitics before exploring changes in global energy demand, the uneven distribution of resources, energy security, contemporary conflicts and the growing role of renewable energy. India is used as the main case study due to its rising energy demand, strategic importance and efforts to balance economic growth with greater energy security.

The research finds that energy continues to influence state behaviour and the balance of power, but the focus of competition is changing. Rather than reducing geopolitical rivalry, the energy transition has expanded it into areas such as critical mineral supply chains and advanced manufacturing, while fossil fuels continue to underpin the global economy. The paper concludes that energy will remain a defining influence on international relations, with strategic advantage increasingly determined by how effectively states adapt to a changing energy landscape instead of simply how many resources they possess.

Introduction

Energy is not simply a factor of production in the global economy; it is the condition that makes industrial life possible. Every major phase of modern development has been shaped by changes in how energy is sourced, transported and controlled. Coal enabled early industrial expansion and territorial concentration of industry. Oil extended mobility and military reach across oceans. Today, the world is entering a more complex phase in which electricity systems and mineral-intensive technologies coexist rather than replace one another. This overlap has made energy systems more politically sensitive and strategically significant than at any earlier point.

The distribution of energy resources across the international system remains deeply uneven, but the consequences of that unevenness have intensified. Production is concentrated in a limited number of regions, while consumption growth is increasingly driven by large emerging economies. This mismatch creates a structural condition in which most states are tied into external supply chains for at least part of their energy needs. Dependence is not uniform, but it is

widespread enough that fluctuations in a few producing regions can generate global economic and political effects. In practice, this means that energy access is no longer only a matter of market efficiency; it is also a question of exposure to external constraints.

That exposure becomes more visible during periods of disruption. Energy markets respond quickly to geopolitical shocks, and those shocks tend to travel through multiple channels at once. A disruption in one region can affect prices, shipping routes, industrial output, and fiscal stability elsewhere. This is why energy has repeatedly appeared at the centre of international crises, even when it is not the original cause of conflict. The strategic significance of infrastructure such as pipelines, refineries and maritime routes reflects this reality. Control over energy flows has become a way of influencing state behaviour without direct territorial control.

These dynamics are not confined to historical cases. In recent years, energy has increasingly been integrated into wider geopolitical strategies through sanctions, export controls and targeted disruptions of infrastructure. At the same time, vulnerability in global supply chains has become more visible as states experience repeated shocks linked to conflict and trade fragmentation. As a result, energy security is now being treated less as a question of price stability and more as an issue of systemic resilience. States are no longer only asking whether energy is available, but whether it remains accessible under stress.

Alongside this continued importance of fossil fuels, the global energy system is undergoing a structural transition. Renewable energy, nuclear power and electrification are expanding rapidly, driven by technological change and long-term policy commitments. However, this shift does not reduce geopolitical competition; it reconfigures it. The focus of strategic attention is moving away from oil production alone toward the control of materials, manufacturing capacity, and electricity infrastructure. This introduces new forms of dependence, particularly in relation to the processing and refining of critical minerals used in clean energy technologies.

The transition is uneven and incomplete. Fossil fuels continue to dominate primary energy supply in most regions, particularly in fast-growing economies where demand is rising faster than infrastructure can adapt. As a result, energy systems are becoming layered rather than sequential. New technologies are being added to existing systems rather than fully replacing them. This creates a period in which states must manage long-term decarbonisation goals while still relying on conventional fuels for stability. The result is not a transition away from geopolitics, but an expansion of it into new domains.

This study examines how these shifts in global energy systems shape international relations. It focuses on how changes in energy production, distribution and technology influence patterns of dependence, competition and strategic behaviour among states. Energy is treated not as a background economic variable, but as a structural force that helps define the boundaries of state power and vulnerability.

Traditional international relations approaches tend to concentrate on oil and gas as instruments of state power, while much of the newer transition-focused work emphasises climate policy and technological change. This division overlooks how deeply interconnected these

systems have become. Fossil fuels and low-carbon technologies now operate within the same global energy structure, often reinforcing rather than replacing each other in the short to medium term. Understanding energy security today requires analysing both simultaneously.

Methodologically, the research uses a qualitative design based on structured literature review and secondary data analysis. It draws on institutional reports, policy documents, and academic work from major international energy organisations and research institutions. Quantitative data is used to support interpretation rather than to test formal models. The study also incorporates comparative case analysis, focusing on emerging economies where energy demand growth, import dependence, and transition pressures intersect most sharply.

Overall, the paper approaches global energy as a dynamic system shaped by technological change, geopolitical competition and structural economic imbalance. It argues that energy security can no longer be understood through supply adequacy alone. Instead, it depends on how states manage exposure to disruption across interconnected systems of trade, infrastructure, and technology. In this sense, global energy is not just a sector of international relations, but one of its core organising forces.

Methodology

This research paper adopts an interdisciplinary qualitative approach, combining both literature review and secondary data analysis to examine the relationship between global energy, geopolitics and international relations in the contemporary international system. This study seeks to understand how dependance, access and control over conventional energy resources influence state behavior and strategic decision making.

The first stage of the research consists of a structured literature review of academic journals, policy papers, institutional reports, articles and scholarly publications related to global energy systems, energy security, geopolitics and international relations. This review will establish the theoretical foundations of the study and identify key debates concerning the implications concerning energy security and geopolitical influence. The research will further employ qualitative policy analysis to examine how global stakeholders conceptualise energy security and respond to perpetually evolving global energy challenges. Key sources will include reports from the International Energy Agency (IEA), Energy Information Administration (EIA), national energy ministries, international organisations and leading policy think tanks. The study will also incorporate secondary quantitative data from institutional data sets and energy reports, including statistics related to energy production and consumption, trade, resource distribution, import dependence and energy demand. These data will be used to support the broader geopolitical analysis rather than for advanced statistical modelling.

To connect broader theoretical discussions with contemporary realities, the study will employ a case study approach. India will serve as the primary case study due to its rapidly growing energy demand, and its increasing strategic importance in the rapidly changing global energy markets. Additional international examples, including the Russia-Ukraine conflict and

tensions in significant energy producing regions, will be referenced where relevant to highlight broader geopolitical trends and tensions as well as energy security challenges.

This research relies exclusively on publicly available academic, institutional and policy sources and does not involve primary data collection. Evidence gathered through literature review and secondary data analysis will be synthesised to examine the relationship between global energy and international relations.

Energy Demand and the Global Distribution of Resources

A defining feature of the geopolitical landscape today is the increasingly large divide between regions that are energy producers and centres of economic consumption. These imbalances can lead to tensions and changes in the priorities of states, as countries compete to secure reliable energy supplies.

Global Trends in Energy Consumption

Global energy demand reached a record high of 655 exajoules (EJ) in 2025, which was an increase of 1.3% from the previous year in 2024 (IEA 43). Overall consumption rates continued to rise even though this was lower than the 2% growth recorded in 2024 (IEA 6). While demand has started to become more stable in many advanced economies, such as the European Union recording a 0.1% decline in 2025 (IEA 44; Energy Institute 6), these reductions have been replaced by growing consumption in emerging economies. At the same time, the pace at which fuel sources are transitioning to renewables has sped up, with solar PV contributing 27% of global energy demand growth in 2025, the largest share of any energy source, ahead of natural gas at 17% (IEA 7). Despite this, fossil fuels still accounted for 87% of the global energy mix at the start of 2025 (Energy Institute 5), suggesting that renewable energy is being added to existing systems rather than just replacing them. This creates a difficult situation for governments, which must continue to secure reliable supplies of oil, coal, and natural gas while also investing in cleaner technologies and the materials that are needed to support them. This challenge can be seen in energy efficiency trends where the IEA estimates that global energy intensity must improve by around 4% annually to meet climate targets, yet improvements reached only 2% in 2025 after several years of slow progress (IEA 11). Recent events have also shown how quickly energy security concerns can outweigh environmental goals. In 2025, emissions in advanced economies rose by 0.5%, compared with 0.3% in emerging markets, the first time this had occurred since the 1990s (IEA 6), as a colder winter and higher gas prices led several countries to increase their use of coal-fired power.

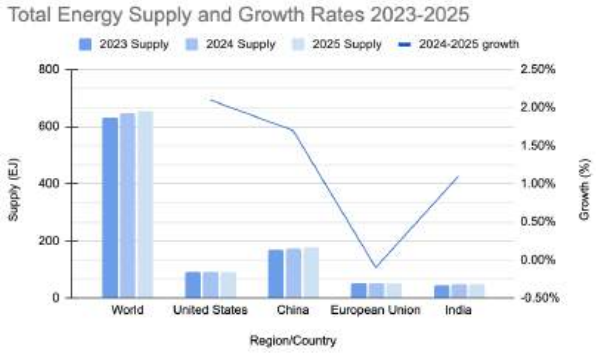


Figure 2: Data synthesised from IEA Global Energy Review 2026

China and India as Drivers of Global Energy Demand

By the start of 2025, the Asia-Pacific region accounted for 47% of global energy demand (Energy Institute 6), making it the largest centre of energy consumption in the world. This shift has important geopolitical implications. As demand growth significantly comes from Asia, major suppliers in the Middle East and Russia have strengthened their focus on Asian markets, particularly countries such as China and India, whose foreign policies continue to be shaped by a strategy of multi-alignment, often referred to as swing states.

China continues to be the world's largest energy consumer, however, its energy profile reflects two differing trends. In 2025, China's energy demand growth slowed to 1.7%, which is down from 2.7% in 2024 as the nation had to face a real estate crisis and shifting industrial priorities (IEA 8). However, China simultaneously accounts for 56% of global renewable energy deployment while its coal demand remains higher than the rest of the world combined as well as accounting for more than 25% of all global energy investment (IEA 8,13). From an international relations perspective, China's drive for energy efficiency and its achievement of 2030 wind and solar targets six years early is an approach focused on greater strategic autonomy, aimed at reducing its vulnerability to potential maritime blockades or global shocks that could cut off 90% of its oil imports (IEA 8,70).

India has emerged as the primary driver of incremental global demand with it projected to account for 25% of the world's increase in energy consumption through 2045 (Federation of Indian Petroleum Industry (FIPI) 41). India's GDP growth of 6.5% is supported by an energy mix that remains 80% dependent on coal, oil and biomass (PwC India 5). However, India faces a core diplomatic concern of structural vulnerability to an extreme extent as it imports about 89% of its crude oil requirements (Sahai 1), leaving its domestic inflation and political stability at the mercy of the 'Middle East risk premium' and volatile international markets.

Energy Resources and Geographic Imbalances

The uneven distribution of energy reserves has created a system of mutual dependence between exporting and importing states, but one that can be easily disrupted and leave states

vulnerable. Global gas production reached 4,124 billion cubic meters (bcm) in 2024 (Energy Institute 35), but it is dominated by just 4 states: the US, Russia, Iran and China. These major providers collectively control 53% of global output (Energy Institute 35). The US accounted for 25% of global gas production in 2024, making it the largest producer in the world (Energy Institute 36). In contrast, gas production in Europe continued to fall, leading to more reliance on imported LNG, mostly from the United States. The Middle East remains an important energy-producing region, which provides 30% of global oil production and 17% of natural gas production (IEA 207). The importance of these supplies was shown during the 2026 Iran conflict. While countries may have the resources needed to enter a conflict, keeping energy supplies flowing during conflicts is challenging. Any disruption to production or transport routes can have economic and political consequences, even far beyond the region itself.

Maritime Chokepoints and the 2026 Systemic Shock

In the era of globalized trade, the physical security of energy is tied to a few irreplaceable chokepoints. The Strait of Hormuz, which transits one-fifth of the world's oil and LNG, has once again become the ultimate pressure point of the international system. The 2026 conflict between the US-Israel and Iran forces led to the effective closure of the Strait of Hormuz, strapping roughly 1, 500 vessels filled with crude oil, LNG and critical fertilizers (Center on Global Energy Policy). The blockade caused sharp price increases and affected around 50% more LNG volume than the 2022 Russian gas crisis (Mahesh). Diesel prices in the United States rose by 42%, while gas prices in Europe increased by 50% (Mahesh). The conflict also showcased how vulnerable energy infrastructure can be during periods of instability. Qatar's Ras Laffan complex facility and the UAE's nuclear power plants became a growing security concern (Mahesh; Center on Global Energy Policy (CGEP)). India was particularly affected as it relies heavily on imported energy. More than 1.6 million tonnes of crude oil and 200,000 tonnes of LNG were stranded on Indian-flagged vessels during the blockade, causing the government to invoke the Essential Commodities Act and prioritise gas supplies for essential sectors (Bakshi 3).

Fragmentation of Global Markets

The uneven distribution of energy resources in addition to growing concerns over the security of trade routes, has contributed to changes in global energy trade. Russian pipeline gas exports to the European Union fell from 155 bcm in 2021 to less than 20 bcm by 2024 (Nawab 1572) following Russia's invasion of Ukraine in 2022 (Mearsheimer). As a result, European countries increased their imports of LNG, particularly from the United States, whose LNG exports to Europe rose by 140% (Nawab 1572). At the same time, Russia expanded energy exports to Asia. The Power of Siberia pipeline was transporting 28 bcm of gas annually by 2025, helping Russia maintain access to foreign markets even after the West imposed sanctions on them (Nawab 1572). These changes have strengthened energy ties between Russia and China, with China benefiting from access to discounted Russian oil (IEA 8). Africa has experienced a different set of challenges. Even though it accounts for around 20% of the world's population, the

continent had the smallest increase in energy demand in 2024, at just 0.2 EJ (IEA 43). Higher energy and fertilizer prices have placed additional pressure on many African economies. At the same time, the continent has become increasingly important in the supply of critical minerals, with countries such as South Africa and Gabon accounting for 57% of global manganese production (Energy Institute 35).

As global energy demand grows continuously, the ways which resources are being used politically and economically are also changing. The 20th century preoccupation with fuel-intensive security is being supplemented by a material-intensive worldview. The supply chains for these minerals are even more concentrated than those of fossil fuels. Therefore, the global distribution of energy demand and resources is no longer a single pillar geared towards Western markers. More so, it is a multipolar, fragmented system where the Asia-Pacific region dictates demand growth, the Middle East remains a volatile supply anchor and the quest for critical minerals defines the new frontier of international relations. For many countries, having sufficient energy supplies now depends as much on transport routes and infrastructure as on the resources themselves.

Energy Security and Geopolitical Competition

As energy systems have become more interconnected, they have also become more vulnerable. Infrastructure that once served mainly commercially now plays an important role in strategic competition and conflict. As the global system enters a disorderly transition, the traditional realism of fossil fuel competition is being overlaid which creates new forms of strategic vulnerability and technological dependency. This section explores how natural gas and transition minerals continue to influence relations between states, as well as focusing on how they shape involvement internationally and affect states' sovereignty and ability to control their own energy futures.

Traditional approaches often focus on the four A's of energy security (availability, accessibility, affordability, and acceptability) but events recently have shown that just using these alone are not enough. The disruptions of 2025 and 2026 highlighted the importance of resilience across supply chains, transport networks, financial systems, and energy infrastructure. In other words, countries are more concerned not only with securing access to energy, but also with making sure that their internal energy systems can continue to function during periods of crisis. This changing approach is shown in global investment trends. According to the IEA, global energy investment is expected to reach USD 3.4 trillion in 2026, with energy security concerns playing an important role in investment decisions (IEA 3).

Much of existing literature on energy and international relations focuses mainly on oil, while often paying less attention to the ways in which natural gas can influence state behaviour. Oil is traded through a global market, however, natural gas depends a great amount on infrastructure such as pipelines and LNG terminals that are very expensive. This creates long-term links between producers, transit states, and consumers, making it more difficult for countries to change their suppliers or export routes in a short amount of time. Research has

identified two main ways in which natural gas can contribute to competition internationally. First, states may want to weaken rival producers in order to protect or expand their own position in the market. One example is Russia's annexation of Crimea, which gave it more control over offshore gas reserves while restricting Ukraine's ability to develop them independently. Second, both exporters and importers place significant importance on transit routes, as disruptions to these can pose a threat to energy supplies and export revenues. As a result, outside powers such as Russia, Qatar, and Iran have oftentimes taken a strong interest in conflicts that involve states located along potential pipeline corridors.

Import dependence can create a situation of coercive vulnerability in emerging "swing states" that are forced to use their military and diplomacy to safeguard their energy supply lines. As an emerging power, India imports 89% of its crude oil (CGEP). Of this 40% (around 2.7m b/d) of oil flows through the vulnerable Strait of Hormuz, the 2026 chokepoints in the global energy map published by the CEIP revealed (Sahai 4). An example of how the world's largest number of oil tankers, that of India, got 'trapped' in the crisis of 2026 and got left stranded at various ports including those of vessels that were flying the Indian flag with a total of 1.6m tonnes of oil and 200,000 tonnes of LNG (Bakshi 3).

Established importers on the other hand have managed to create strategic space for themselves by building up a network of inland infrastructure to bring in oil from countries such as Russia and Central Asia via pipelines. China for instance is also working to mitigate the 'Malacca Dilemma' which refers to the risk of a naval blockade which could strand 90% of China's seaborne oil imports (San-Akca et al. 2). To achieve this end, large oil companies of China have been positioned as strategic storage partners. They have thus taken significant volumes of oil off the global market and stockpiled them on their onshore facilities.

The capacity to withstand crisis is increasingly dependent upon a country's Strategic Petroleum Reserves (SPRs), thus there is an expanding 'resilience gap' among major importers. A look at preparedness of key regional states in the event of a 2026 Iran conflict, for example, immediately reveals some stark differences between Japan, China, and India, and between established and emerging 'swing states': Japan has 254 days of demand cover in its SPR, the highest in the world (S&P Global). China has stockpiled around 100 days of onshore crude oil in Strategic Petroleum Reserves (S&P Global). India's petroleum reserve of 75 days of demand of 1.7 million barrels of crude as of 2026 is not sufficient to insulate the country from the 'Middle East risk premium' (Sahai 10, S&P Global). Thus, domestic inflation and political stability would be severely affected. Similar concerns however are now beginning to emerge in the United States, given that its Strategic Petroleum Reserve has been drained to record lows in 2026, in a bid to capitalize on higher price caused by the conflict in Ukraine as well as the ongoing fighting in Iran, in order to provide around 125 days of cover against an unexpected disruption to import requirements (Mahesh).

Physical transit security remains concentrated in irreplaceable chokepoints, rendering the global order geographically brittle. The Strait of Hormuz, facilitating one-fifth of the world's oil and LNG has once again become the ultimate pressure point of the international system. The

2026 blockade trapped roughly 1,500 vessels and disrupted 20% of global supply, a shock 50% larger than the 2022 Russian gas crisis (CGEP, Mahesh). While states have attempted to respond through naval initiatives like ‘Project Freedom’ by the USA, the asymmetric use of drone swarms and anti-ship missiles has challenged the ability of traditional hard power to enforce maritime security. Consequently, many countries are looking towards accelerated electrification as a means of converting imported fuel risk into domestic control. This shift has tangibly improved energy security in major importing regions. For example, investments in renewables and nuclear power avoided around 260 billion USD in fossil fuel import costs globally in 2025 (IEA 10).

Energy infrastructure has morphed from a pathway to independence into a battlefield instrument and a lever of coercive diplomacy. For example, Russia has successfully mitigated Western sanctions by redirecting its energy rents to Asia through the Power of Siberia pipeline. As well as this, Washington's 2026 sanctions on Hengli Petrochemical prompted Beijing to deploy unprecedented domestic blocking rules to counter US jurisdictional reach. This signals a new era where energy-driven trade warfare and legal warfare are primary tools of statecraft. Beijing’s recent restriction on the export of materials like gallium, germanium and antimony to the United States demonstrates a new form of an energy weapon used for diplomatic leverage. China’s major concentration of rare earth elements (REEs), has prompted US-led initiatives to restore supply chains, however, upfront capital costs in diversified regions are much higher than those in traditional energy-producing regions.

Contemporary Conflicts and Global Energy Politics

In the study of international relations (IR), the longstanding liberal institutionalist premise of ‘Wandel dutch Handel’ (Change through Trade), which has suggested that economic entanglement, particularly in energy, would serve as a safeguard against conflict, has been definitively replaced by a paradigm of leveraged independence.

The Russia-Ukraine conflict serves as the definitive case study for the collapse of the 20th century energy order. Historically, the relationship between Moscow and Kyiv was defined by a volatile symbiosis rooted in Soviet-era infrastructure. Ukraine inherited a vast network, including the Brotherhood and Soyuz pipelines, which once transited 80% of Russia’s piped gas exports to Europe. By the late 2010s, the United States had become a major exporter of liquefied natural gas (LNG) following the shale gas revolution. As US LNG exports expanded, Washington sought a larger share of the European gas market, supporting efforts to diversify Europe's energy supplies and reduce its dependence on Russian pipeline gas. This intensified competition in the European gas market and added an important economic dimension to the broader geopolitical rivalry between Russia and the West. Russia utilized its vertical monopoly over these networks as a tool of coercive diplomacy long before the 2022 invasion. By February 2022, this evolved into a full-scale strategy of energy weaponization, evidenced by Russia’s strategic withholding of 5.5bcm of gas in late 2021 to pressure the certification of the Nord Stream 2 (Nawab 1579).

At the same time, growing US LNG exports and efforts to diversify Europe's gas supplies reduced reliance on Russian pipeline gas, threatening Russia's dominant position in the European energy market and adding an important energy dimension to the wider geopolitical tensions surrounding the conflict. The conflict has catalyzed a decolonization of the European energy map as Russian pipeline gas flows to the EU plummeted by 70% from 155 bcm in 2021 to under 20bcm by 2024. This decoupling was physically manifested in the September 2022 Nord Stream sabotage which eliminated 110bcm of potential capacity (Nawab 1580). Consequently, Ukraine's role as a transit hub has undergone systemic deterioration, with volumes falling from a pre-2019 average of 90bcm to just 15 bcm in 2024. (Nawab 1579). In response, Moscow has demonstrated adaptive statecraft, redirecting energy rents to Asia. The Power of Siberia pipeline doubled its exports to 38 bcm annually to China by 2025 (Nawab 1572), illustrating a reconfiguration of the global balance of power where Moscow trades European dependency for an asymmetric partnership with Beijing.

The Middle East still is the world's "central bank of oil" with over 40% of proven oil reserves and the only spare capacity to speak of (Energy Institute 5) . Yet, 2025-2026 will be a more volatile year in terms of geopolitics as well as energy organization in the region than previous years. A significant chapter in the region's energy politics is that of the United Arab Emirates' (UAE) landmark withdrawal from the Organization of Petroleum Exporting Countries (OPEC), which took place on May 1, 2026. The exit of the UAE from OPEC marks a point of significant divergence in policy among a group of producers who are increasingly bound by the restrictions of production targets versus a growing group of countries seeking to capitalize on their resources in order to kick start the economies of their countries. In light of the UAE's withdrawal from the organization, Abu Dhabi National Oil Company (ADNOC) formally outlined a 5-year, $150 billion capital expenditure program as it transitioned towards an autonomous 'energy state' (IEA 33). Meanwhile, the OPEC+ group will continue to grapple with highly volatile oil markets as well as an increasingly down-trending global demand for energy. In August 2025, for example, OPEC+ reversed 80% of the 2+ year-old production reductions that were implemented in 2023 in order to raise production by the largest amount in the associated history in order to defend market share in the face of significant increases in non-OPEC+ supply, especially from the Americas and, for the first time in decades, from Latin America (FIPI 2).

Meanwhile, an "energy risk premium" has been driving up the prices of producers' services in the Middle East and elsewhere, further straining the related fiscal policies of major energy producers. Saudi Arabia's fiscal policy is, for example, currently based on a "break-even price" of $88 per barrel of oil, up sharply from $68 per barrel in 2012, as the kingdom has provided huge public transfers in recent years in order to maintain the stability of its population (Cornell 21). While the actual cash cost of pumping a barrel of oil out of the ground in Saudi Arabia is extremely low (around $9 per barrel), the fiscal break-even price includes public sector wages, social subsidies, defense, and massive national infrastructure projects.. Such a high price in turn has the potential to become self-fulfilling as producers may be forced to maintain current oil prices in order to preserve their rule in the face of declining oil prices and high costs of

production in the meantime, at the same time that those same prices are encouraging development of alternative sources of energy.

Changes in energy markets and history of producers in the region are unleashing a series of transformations which are taking Middle East producers out of their historical dependencies towards Western markets. Asia now accounts for 47% of total global energy demand (Energy Institute 6). This has led to a strategic decoupling, for example, China has become Saudi Arabia's primary customer by a significant margin. This shift undermines the traditional oil-for-security logic that linked the region to the United States, as producers now prioritise the economic health of Asian emerging markets over Western security alliances.

While Eurasia transformed pipeline politics, the 2026 conflict in the Middle East, frequently referred to as the 'third Gulf War', re-established the enduring vulnerability of maritime energy transit. Triggered by US-Israeli strikes on February 28, 2026, the conflict led to the immediate closure of the Strait of Hormuz. Iran's closure of the Strait was a calculated disruption aimed at the global commons. In a novel move, Iran attempted to institutionalise its control by levying tolls on passing vessels and utilising GPS disruption and drone swarms to prohibit tankers. Strikes on Iranian South Pars gas fields were met with retaliatory drone barrages on Qatar's Ras Laffan LNG complex and the UAE's nuclear power plants. Notably, the damage to Ras Laffan resulted in the loss of 17bm/yr of LNG capacity, which will remain off the market even if the Strait reopens (IEA 33).

The persistence of energy-driven conflict has forced states and international organisations to move from 'just-in-time' market reliance to 'just-in-case' strategic independence'. The EU's success in reaching a 42.5% renewable electricity share by Q1 2025 (Nawab 1572) is a direct byproduct of war-driven necessity. By 2025, 33% of Russian oil exports successfully bypassed G7 price caps via a 'shadow fleet' of aging tankers with obscured ownership (Nawab 1572). This has forced the West to adopt the use of competing legal frameworks and secondary sanctions as a substitute for traditional force.

Contemporary conflicts prove that energy is not a mere backdrop to international relations but the primary enabler of state power, The combination of pipeline decoupling in Europe, the fracture of OPEC unity with the UAE's exit and the institutionalization of chokepoint control in the Middle East has definitively ended the era of globalized energy stability. These conflicts have accelerated the fragmentation of global energy markets into regionalised blocs, ensuring that energy remains a perennial vector of instability.

The Emerging Role of Renewable Energy

The transition to low-carbon systems is no longer viewed solely through the lens of climate mitigation but as a decisive instrument of strategic autonomy. As the global energy system enters the age of electricity, the merging geopolitics of material-intensive supply chains and technological sovereignty continue to overlay fossil-fuel realism. In 2025, for the first time on record, a modern renewable source, Solar PV, contributed the largest single share of global energy demand growth, signalling that the structural shift away from hydrocarbon dominance has

reached a tipping point (IEA 4). Far from eliminating dependency, the renewable energy transition is transforming its nature. As reliance on continuously traded fuels gives way to dependence on the materials embedded in clean-energy technologies, new questions emerge regarding energy resilience and the future structure of international relations.

The global adoption of renewable energy has defied economic uncertainty and high interest rates, reaching record-breaking expansion levels in 2025. Annual renewable capacity additions rose to a record 800 gigawatts (GW) in 2025, a 16% increase over the previous year, with solar PV alone accounting for 75% of these additions (IEA 31). This explosive growth is driven not only by cost competitiveness but by transition securitisation, where states view domestic generation as a bulwark against the volatility of international fuel markets.

China continues to set the global pace, commissioning nearly 370 GW of solar PV and 117 GW of wind in 2025 alone, which is more than one quarter of all global energy investment (IEA 32, 6). However, the 2026 landscape shows a significant surge in adoption within emerging markets and developing economies (EMDE) heavily affected by the 2026 Hormuz crisis. For example, Pakistan's continuing influx of Chinese solar panels which is ongoing into 2026 to insulate its economy from expensive grid-connected LNG generation (IEA 8).

From the perspective of international relations, the most profound implication of renewable energy is its capacity to reverse the trend of foreign energy reliance that has dictated state behavior since the Industrial Revolution. Unlike fossil fuels, which require a continuous and recurring flow of imported commodities, renewable energy allows states to convert a one-time embodied dependency (the procurement of panels or turbines) into a long-term stream of domestic energy sovereignty.

Investments in renewables and nuclear power across the five major importing regions (China, EU, Japan/Korea, Southeast Asia and India) avoided an estimated 260 billion USD in fossil fuel import costs in 2025 (IEA 10). For states like India, which imports the majority of its crude oil, the deployment of 500GW of non-fossil fuel capacity by 2030 is not merely an environmental goal but a strategic hedge against maritime chokepoints (Joshi, IEA 10). In the first quarter of 2026, 15 African countries reported record-high solar imports exceeding 400 million USD, as households and businesses sought to insulate themselves from the Hormuz conflict (IEA 8). As developed countries maximise 'homegrown' resources, such as wind in North Sea or solar in the Iberian Peninsula, the historical necessity for military projection to secure distant oil fields begins to dissipate, potentially decoupling energy from traditional security concerns.

Green hydrogen, produced via the electrolysis of water using renewable electricity, has emerged as the vital missing piece for decarbonizing hard-to-abate heavy industries such as steel, chemicals and shipping. While still in its early stages, low-emissions hydrogen investment is expected to rise 50% to nearly 10 billion USD in 2026, with Europe, the US and China leading the pipeline (IEA 25). From a geopolitical perspective, green hydrogen offers the potential for new energy alliances. Countries with vast renewable potential but low domestic demand (e.g Chile, Morocco or Australia) are positioning themselves as 'Hydrogen Superpowers', aiming to

replace OPEC states as the primary exporters of clean energy. However, this hydrogen economy faces significant technological constraints, including uncertainty over electrolyser dominance (Alkaline vs. PEM) and high capital costs that have caused some investors to adopt a wait-and-see approach.

While renewables reduce dependence on fossil fuels, they increase dependence exponentially on critical minerals and specialised technology. This materiality paradox suggests that the transition is not an escape from geopolitics but a relocation of competition to the back-end of the supply chain. This shift has created a high geographical concentration that mirrors or exceeds that of OPEC. This concentration makes the clean energy transition vulnerable to mineral weaponisation as seen in 2025 when export restrictions on gallium and germanium impacted Western technology sectors.

The variable nature of wind and solar necessitates massive investments in Battery Energy Storage Systems (BESS) and firm dispatchable power to maintain grid stability. In 2025, BESS was the fastest growing power technology, with capacity additions rising by 40% to reach almost 110 GW (IEA 5). Despite this, the IEA warns that until long duration storage (LDES) matures, natural gas and nuclear will remain essential for 'firmness', potentially creating a prolonged hydrocarbon dependence through 2035 (McKinsey & Co., IEA). The most significant bottleneck in 2026 is not generation but transmission infrastructure. Procurement duration times for large power transformers have doubled to 4 years while specialised direct-current (HVDC) systems now have wait times exceeding 5 years (IEA 197). Without a modern and resilient grid, states could risk a connection crisis that could stall the integration of new renewable capacity.

A fresh insight from the 2026 crisis is the role of distributed energy resources (DERs) in enhancing national resilience. Unlike centralised power plants, which are vulnerable to infrastructure targeting (as seen in the 50% grid loss in Ukraine), decentralized rooftop solar and microgrids provide a form of citizen-led energy security (Nawab 1572). In India the PM Surya Ghar: Muft Bill Yojana has launched a rooftop revolution, making solar energy accessible to 10 million households (FIPI 115). This empowers citizens as 'energy owners' and reduces the systemic risk of a single point of failure in the national power grid. Similarly, the rapid uptake of heat pumps in Europe (up 17% in Q1 2026), represents a structural shift where consumers are using electrification as a strategic hedge against volatile gas prices (IEA 299).

The emerging role of renewable energy has transformed energy security from a technical logistics problem to a core pillar of national survival and industrial strategy. While the transition offers a pathway to decouple economic growth from fuel import risks, it simultaneously entrenchies a new set of technological and mineral dependencies that prioritize those states capable of dominating the chain. In 2026, the international system has seen record clean energy investments (2.2 trillion US) coexist with a resurgence in coal and gas power to meet the skyrocketing electricity demand from artificial intelligence and data centres (IEA 297, 302). The strategic focus is no longer on 'if' the world will transition, but on who will control the infrastructure and materials of the new era. For 'swing states' like India and the ASEAN bloc, the challenge lies in leveraging their renewable ambitions to achieve Strategic Autonomy while

navigating the supply chain bottlenecks and mineral monopolies that define the 21st century multipolar order.

India Case Study

India's rapid economic growth and rising energy demand have made it an increasingly important player in global energy markets. S&P Global describes India as a "stabilizing factor" in the global energy system. As demand shifts towards the Indo-Pacific, India's role in global energy markets is expected to grow further. India is projected to account for nearly 25% of global energy demand growth by 2045 (FIPI 83). Meeting this demand has required India to build relationships with a wide range of energy suppliers while reducing the risks associated with import dependence. India's experience shows how an emerging economy can balance energy security, economic growth, and the transition to cleaner energy.

India is currently the world's third-largest energy-consuming economy, driven by rapid urbanisation and the growth of energy-intensive industries such as data centres and AI-related infrastructure (PwC India 5). In 2025, the country's total energy supply reached approximately 49 exajoules (EJ), an increase of 1.1% from the previous year, despite a record monsoon that reduced demand for cooling and agricultural pumping (IEA 39). At the same time, electricity is becoming an increasingly important part of India's energy system. Investment in the power sector has grown significantly and now accounts for almost half of the country's total energy spending (IEA 440). India's economy grew by 6.5% in 2025, but energy consumption did not increase at the same pace, reflecting improvements in energy efficiency (FIPI 78). These gains have helped moderate growth in energy demand at a time when pressure on infrastructure continues to increase. India has also become the world's largest recipient of development finance institution (DFI) funding for clean energy, receiving USD 2.4 billion in 2024 alone (IEA 244). This investment has strengthened India's role in global energy markets. According to the Ministry of External Affairs, stable energy prices are a shared global interest, and India has sought to support this through its engagement with both producers and consumers.

To reduce its dependence on imported energy, India has increased efforts to develop domestic oil and gas resources. The Petroleum and Natural Gas (PNG) Rules 2025 introduced a stabilization clause aimed at protecting investors from major legal or fiscal changes, making exploration projects more attractive to foreign companies (FIPI 89). The government has also set a target of expanding exploration acreage to 1 million sq km by 2030 and has reduced offshore "No-Go" areas by 99% (Sahai 4). One area of focus has been the Andaman Basin, a deepwater frontier that has attracted renewed exploration because of its geological potential. Although the ANDW-7 exploratory well in early 2026 confirmed an active thermogenic petroleum system, subsequent drilling has produced more limited results, highlighting the uncertainty and long timeframes associated with frontier exploration. Rather than providing an immediate solution to import dependence, the basin illustrates India's broader effort to diversify domestic energy sources despite significant geological and commercial risks. These efforts have also been

supported by the National Data Repository, a cloud-based platform that provides geoscientific data to exploration companies and international partners (FIPI 89).

India aims to increase the share of natural gas in its primary energy mix to 15% by 2030 as part of its goal of becoming a more gas-based economy (PwC India 6). Supporting this objective is the continued expansion of the National Gas Grid, which now includes more than 24,945 km of operational pipelines (PwC India 251). Despite this progress, challenges remain. Although India is the world's fourth-largest LNG importer, many regasification terminals have operated at less than 30% of their capacity due to inadequate downstream infrastructure and connectivity (PwC India 257). In response, the LNG Terminal Regulations 2025 require developers to submit evacuation plans and provide bank guarantees to ensure that new projects are linked to actual demand. One important initiative is the development of LNG corridors, which aim to shift a significant share of long-distance freight transport from diesel to natural gas. Government projections suggest that this could lead to around 400,000 LNG-powered trucks on Indian roads by 2040, creating a substantial new source of demand while reducing reliance on diesel imports (PwC India 274). At the same time, policymakers are considering the creation of a strategic gas reserve with a capacity of 4 billion cubic metres (PwC India 20). By using depleted reservoirs for storage, such a reserve could help cushion the impact of future supply disruptions.

Coal continues to play a central role in India's energy system. Coal-fired power plants still generate around 71% of the country's electricity, and investment in coal supply is expected to reach USD 13 billion in 2026 as India works towards its target of producing 1.5 billion tonnes of coal annually by 2030 (IEA 28, 233). The country has also increased spending on coal transport infrastructure, with investment reaching USD 7 billion to support domestic production and reduce dependence on imported coal (IEA 42). At the same time, India is expanding its nuclear energy programme as part of its long-term effort to diversify electricity generation. A major reform in 2025 ended the state monopoly on nuclear power and allowed private and foreign companies to build and operate reactors, with foreign ownership permitted up to 49% (IEA 235). The government aims to reach 100 GW of nuclear capacity by 2047, with Small Modular Reactors (SMRs) expected to play an important role in providing reliable electricity alongside renewable energy sources (IEA 235). Demand for nuclear power is also growing among technology companies, which accounted for 10% of new nuclear power purchase agreements as they seek stable, low-emission electricity for data centres (IEA 83).

India's expansion of renewable energy is driven not only by environmental concerns but also by economic and strategic considerations. In 2024, the growth of renewable energy helped the country avoid nearly USD 14.9 billion in fossil fuel imports (FIPI 58). Investment in solar PV reached USD 20 billion in 2025 and accounted for around one-fourth of the growth in India's total energy investment over the previous five years (IEA 233). Supporting this effort is the Production Linked Incentive (PLI) scheme, a USD 3 billion programme designed to expand domestic manufacturing of solar modules and batteries and reduce reliance on overseas supply chains. The transition to cleaner energy, however, brings new challenges. Technologies such as electric vehicles, batteries, and solar panels require large quantities of critical minerals. An

electric vehicle, for example, uses around six times more mineral inputs than a conventional car (IEA 5). To address this, India has sought to strengthen international partnerships and increase investment in domestic research and development. Recent advances have also highlighted the country's technological potential. Researchers at IIT Bombay developed silicon-perovskite tandem solar cells with an efficiency of 29.8%, a development that could improve the performance of future solar technologies and support India's long-term energy goals (FIPI 53).

Energy security also has important economic implications. To help manage the impact of external shocks on the domestic economy, the government has proposed an economic stabilization fund worth INR 573 billion (USD 6.20 billion) (Bakshi 9). The fund is intended to reduce the effects of sudden increases in energy prices, particularly during periods of international disruption. One example was the 2026 crisis, which damaged major export hubs, reduced expectations of an LNG surplus, and forced India to purchase more expensive spot cargoes. The government has also introduced measures to improve the financial position of electricity distribution companies. These include payment security mechanisms and escrow accounts designed to ensure more reliable payments across the power sector. Such reforms are particularly important given that distribution companies owed more than USD 9 billion in unpaid dues in 2025 (IEA 226). By reducing financial uncertainty, policymakers hope to attract greater investment in the energy sector and lower financing costs, which remain around 80% higher than those in advanced economies (IEA 226).

India's position in the global energy system has been strengthened by its efforts to diversify energy supplies and expand domestic capacity. By importing energy from 41 countries and reaching its 50% non-fossil capacity target five years ahead of schedule, India has reduced its exposure to supply disruptions while supporting continued economic growth (S&P Global, FIPI 51). The 2026 Hormuz crisis highlighted the value of this approach. As India works towards its Net Zero 2070 target, India's reliance on a mix of coal, natural gas, nuclear power, and renewable energy offers one possible path for other emerging economies facing similar challenges (FIPI 51).

Future of Global Energy Geopolitics

Global energy markets are changing rapidly, but the transition to cleaner energy has not reduced geopolitical competition. According to the International Energy Agency (IEA), global energy investment is expected to reach USD 3.4 trillion in 2026, with around USD 2.2 trillion directed towards renewables, nuclear power, electricity networks, storage, and other low-carbon technologies (IEA 13). Rather than ending competition over energy, this shift is changing what states compete for. Alongside oil and gas, manufacturing capacity, critical minerals, and electricity infrastructure have become increasingly important sources of economic and strategic influence.

Patterns of investment reflect this shift. Advanced economies and China now account for more than 70% of global energy investment, with China alone responsible for over one-quarter of global spending (IEA 14). At the same time, National Oil Companies (NOCs), particularly in

the Middle East, have expanded their role in upstream oil and gas investment. Their share has increased from less than 10% in 2015 to more than 20% in 2026, showing that hydrocarbons remain central to global energy markets despite rapid investment in low-carbon technologies (IEA). This has also strengthened energy links between Gulf producers and Asian economies, where demand continues to grow more quickly than in Europe or North America.

The transition has also shifted attention towards the production of clean energy technologies. In 2025, China accounted for 75% of global clean-energy manufacturing investment, including 80% of the lithium-ion battery supply chain and 95% of photovoltaic wafer production capacity (IEA 12). Efforts to diversify production face significant financial barriers, as projects outside established manufacturing centres such as China and the Democratic Republic of the Congo require substantially higher upfront investment (IEA 48). Export restrictions on gallium, germanium, and antimony during 2025 further demonstrated how control over mineral processing can influence international trade and supply chains.

Although investment is shifting towards electricity and renewable technologies, fossil fuels continue to play a major role. Global investment in oil, gas, and coal is projected to recover to USD 1.2 trillion in 2026, while final investment decisions for new gas-fired power projects reached their highest level in 25 years during 2025-26 (IEA 5, 10). As renewable generation expands, natural gas is increasingly used to provide reliable electricity when wind and solar output is low, illustrating that the transition is changing the role of hydrocarbons rather than replacing them altogether.

The rapid expansion of artificial intelligence (AI) and digital infrastructure is creating a new source of energy demand with important geopolitical implications. Global investment in data centres reached USD 565 billion in 2025, exceeding investment in oil supply, while the energy infrastructure needed to support them accounted for USD 105 billion of this total (IEA 82). In the United States, orders for gas turbines intended to power data centres now exceed those of any other country (IEA 20, 78). At the same time, growing electricity demand has renewed interest in nuclear power, particularly Small Modular Reactors (SMRs), while venture capital investment in nuclear fusion reached USD 2 billion in 2025 (IEA 153). These developments suggest that future competition will depend not only on access to energy, but also on the ability to provide reliable, low-carbon electricity for emerging technologies.

Governments are increasingly responding through industrial policy. Carbon pricing mechanisms generated an estimated USD 107 billion globally in 2025, providing both incentives to reduce emissions and additional public revenue for investment in clean technologies (IEA). At the same time, policies such as the United States' One Big Beautiful Bill Act (OBBBA) and China's 15th Five-Year Plan place greater emphasis on domestic manufacturing, strategic industries, and supply chain resilience rather than relying solely on international markets (IEA 206, 230). This marks a broader shift away from cost-driven globalisation towards greater state involvement in energy and industrial development.

Recent geopolitical events have reinforced these trends. Disruptions to major shipping routes and attacks on critical energy infrastructure during 2026 highlighted the vulnerability of

global supply chains. As a result, many governments are placing greater emphasis on strategic reserves, domestic production, and more resilient energy systems. Rather than relying entirely on international markets, countries are increasingly seeking to reduce exposure to external shocks while maintaining access to the technologies and materials needed for the energy transition. Energy security has therefore become closely linked to industrial competitiveness, technological capability, and the resilience of national infrastructure.

A key feature of the current energy transition is the growing emphasis on resilience rather than simple supply security. Traditional measures such as import dependence are becoming less useful on their own, as attention shifts towards broader indicators such as financing conditions and the stability of energy systems. According to the IEA, the cost of capital for green energy projects in Emerging Market and Developing Economies (EMDEs) remains around 80% higher than in advanced economies, creating a significant barrier to investment and slowing the pace of the energy transition in many regions (IEA 226). This gap risks reinforcing global inequalities in access to clean energy, even as overall investment increases.

In response, governments and institutions are adopting more diversified approaches to energy security. One approach involves the expansion of Distributed Energy Resources (DERs), particularly rooftop solar and small-scale generation, which allows households and local systems to reduce dependence on centralised infrastructure. Another involves the development of strategic reserves beyond oil, including critical minerals stockpiles and gas storage facilities such as salt caverns, designed to provide buffers during supply disruptions (IEA 5; PwC India 20). These measures reflect a broader shift towards managing volatility rather than eliminating it.

There is also growing interest in circular economy strategies, particularly in relation to critical minerals. Recycling and material recovery are increasingly seen as ways to reduce long-term dependence on concentrated supply chains by turning one-time material use into a more continuous domestic resource stream (IEA). While still limited in scale, such approaches are becoming more relevant as demand for minerals used in batteries, renewable energy systems, and digital infrastructure continues to rise.

Taken together, these developments suggest that energy security is being redefined. It is no longer determined only by access to resources or geopolitical control over supply routes, but also by the ability of states to manage costs, attract investment, and adapt their energy systems to shocks. Countries that are able to combine diversified supply chains, technological capability, and flexible domestic systems are likely to be better positioned in this evolving environment.

Conclusion

This study has examined the relationship between global energy and international relations by tracing the evolution of energy from a largely economic resource to a central determinant of geopolitical power. The evidence presented throughout the paper suggests that energy has never been politically neutral. From the industrial expansion driven by coal, through the oil-dominated order of the twentieth century, to today's competition over critical minerals and electricity infrastructure, changes in energy systems have consistently reshaped patterns of

cooperation, conflict, and state behaviour. While the technologies underpinning global energy continue to evolve, the strategic importance of controlling energy resources, infrastructure and supply chains remains a defining feature of international relations.

The research demonstrates that energy security can no longer be understood solely through the traditional objective of securing reliable fuel supplies. Contemporary energy systems are characterised by deep interdependence and increasing exposure to geopolitical disruption. Recent conflicts have highlighted that vulnerabilities extend far beyond oil production itself. Maritime chokepoints, pipeline networks, electricity grids, LNG infrastructure and the processing of critical minerals have all become strategic assets capable of influencing diplomatic decision-making and economic stability. As a result, resilience has emerged as an equally important dimension of energy security alongside availability and affordability.

The study also shows that the global energy transition is changing rather than reducing geopolitical competition. Renewable energy has expanded rapidly, supported by technological progress and sustained investment, yet this has not diminished the strategic importance of fossil fuels. Instead, the international system is entering a prolonged period in which conventional fuels and low-carbon technologies operate simultaneously. Oil and natural gas continue to underpin industrial production, electricity generation, and military capability in many parts of the world, while renewables increasingly shape investment decisions and industrial policy. Rather than replacing one geopolitical order with another, the transition has created overlapping systems of dependence that require states to manage both legacy fuel security and emerging technological competition at the same time.

Control over critical minerals and clean-energy manufacturing has become one of the clearest examples of this shift. The concentration of processing capacity within a small number of countries demonstrates that new technologies do not eliminate strategic dependence but relocate it. Competition has therefore expanded beyond the ownership of natural resources to include manufacturing capability, technological innovation, and control over supply chains. This represents a significant transformation in the nature of energy geopolitics, where influence increasingly derives from industrial capacity as much as from resource endowment.

The analysis of contemporary conflicts further illustrates that energy continues to shape international relations in direct and indirect ways. The Russia-Ukraine conflict, tensions in the Middle East, and the disruption of major maritime routes demonstrate how energy infrastructure has become both a strategic objective and a political instrument. Sanctions, export restrictions, attacks on infrastructure, and efforts to redirect trade flows all reflect the extent to which energy considerations influence foreign policy decisions. These developments have accelerated the fragmentation of global energy markets and encouraged governments to prioritise diversification, strategic reserves, and domestic production over purely market-based approaches.

India provides an important example of how emerging powers are responding to these changing conditions. Rising energy demand, heavy import dependence, and ambitious decarbonisation targets place India at the centre of many of the challenges explored throughout this study. Rather than relying on a single strategy, India has pursued diversification across

suppliers, expanded domestic exploration, invested heavily in renewable energy, strengthened electricity infrastructure, and continued to develop conventional fuels where necessary. This balanced approach reflects the reality that energy transitions occur gradually and that economic development requires both reliability and flexibility. India's experience suggests that strategic autonomy in the twenty-first century depends less on complete energy independence than on the capacity to absorb external shocks without undermining economic or political stability.

More broadly, the findings indicate that the relationship between global energy and international relations is becoming increasingly multidimensional. Questions of military security, industrial competitiveness, technological innovation, climate policy and economic resilience are now closely interconnected. Decisions made within national energy sectors increasingly influence diplomatic relationships, trade patterns, and broader strategic alignments. Energy policy has therefore become an integral component of foreign policy rather than a separate area of domestic economic management.

Several limitations should also be acknowledged. This study relies primarily on qualitative analysis and secondary data drawn from institutional reports and academic literature. While these sources provide a broad understanding of contemporary developments, future research could incorporate quantitative modelling to examine more precisely the relationship between energy dependence and geopolitical behaviour. In addition, the pace of technological innovation means that many aspects of the global energy transition remain uncertain. Developments in battery storage, hydrogen, nuclear technologies, and critical mineral recycling could substantially alter current patterns of dependence over the coming decades.

Future research should also pay greater attention to the geopolitical implications of artificial intelligence, electricity demand and digital infrastructure. As data centres become major consumers of electricity and governments compete to secure reliable low-carbon power, energy security is likely to become increasingly intertwined with technological leadership. Similarly, greater attention should be given to the role of developing economies beyond China and India, particularly in Africa and Southeast Asia, where rapid demand growth and expanding renewable deployment are likely to reshape future energy markets.

Ultimately, this study argues that the future of international relations will not be determined by the disappearance of energy geopolitics but by its transformation. The competition that once centred on oil fields and pipeline routes is expanding to encompass electricity networks, critical minerals, advanced manufacturing, and technological capability. States that are able to diversify their energy systems, strengthen domestic resilience, and adapt to changing patterns of dependence will be better positioned to navigate an increasingly fragmented international order. Energy has always influenced the distribution of power within the international system. In the decades ahead, it will continue to do so, although through new technologies, new forms of dependence and new geopolitical rivalries that will shape the evolving structure of global politics.

Works Cited

Bakshi, Parul. 2026. 'Risk and Resilience: India's Energy Security in a Volatile Middle East'. Orfonline.Org, OBSERVER RESEARCH FOUNDATION (ORF), March 30. https://www.orfonline.org/research/risk-and-resilience-india-s-energy-security-in-a-volatile-middle-east.

Brew, Gregory. 2026. 'Iran's New Oil Weapon: How America Can Protect Itself—and the Global Economy'. Foreign Affairs, Foreign Affairs Magazine, May 6. https://www.foreignaffairs.com/iran/irans-new-oil-weapon?utm_source=google&utm_medium=cpc&utm_campaign=dsa_middle_east_tfd_am&gad_source=1&gad_campaignid=21047307262&gbraid=0AAAAAD7NbWm6YK-ZmyTKvUIeyCI1A3mJ8&gclid=Cj0KCQjw4orUBhCjARIsAIbF3qz3gwSBHQvpyE8adZlpxSlw6EPIQPZ_-iG-H4mDO105SQQ_AfEHuJoaAmf4EALw_wcB.

Buddeberg, Eva. 2019. 'Is There Moral Progress?'. Analyse & Kritik 41 (2): 195–204. https://doi.org/10.1515/auk-2019-0012.

Caballero-Anthony, Mely, and Collin Koh. 2008. The Way Forward for Energy Security. S. Rajaratnam School of International Studies. JSTOR. https://www.jstor.org/stable/resrep05806.

Cornell, Phillip. 2012. 'Regional and International Energy Security Dynamics: Consequences for NATO's Search for an Energy Security Role'. GCSP Geneva Papers -Research Series 5 (January). https://www.files.ethz.ch/isn/139888/regional%20and%20international%20energy%20security%20dynamics.pdf.

Energy Institute. 2024. 'Statistical Review of World Energy'. In Statistical Review of World Energy. https://www.energyinst.org/statistical-review.

Federation of Indian Petroleum Industry (FIPI). 2025. POLICY and ECONOMIC REPORT OIL & GAS MARKET. https://www.fipi.org.in/Policy_Economic_Report/Policy%20Economic%20Report%20-%20July%202025.pdf.

Hart, David M. 2026. 'Innovation and the Energy Crisis'. In Cfr.Org. Council on Foreign Relations. https://www.cfr.org/articles/innovation-and-the-energy-crisis.

IEA. 2021. 'The Role of Critical Minerals in Clean Energy Transitions – Analysis'. In IEA. https://www.iea.org/reports/the-role-of-critical-minerals-in-clean-energy-transitions.

International Energy Agency (IEA). 2025. 'World Energy Investment 2025 – Analysis - IEA'. In IEA. https://www.iea.org/reports/world-energy-investment-2025.

International Energy Agency (IEA). 2026a. 'Global Energy Review 2026 – Analysis - IEA'. In IEA. https://www.iea.org/reports/global-energy-review-2026.

International Energy Agency (IEA). 2026b. 'World Energy Investment 2026 – Analysis - IEA'. In IEA. https://www.iea.org/reports/world-energy-investment-2026.

'Iran Conflict Brief: How the Iran Standoff Is Rewriting US-China Relations - Center on Global Energy Policy at Columbia University SIPA | CGEP'. 2026. In Center on Global Energy

Policy at Columbia University SIPA | CGEP. https://www.energypolicy.columbia.edu/iran-conflict-brief-how-the-iran-standoff-is-rewriting-us-china-relations/.

Luomi, Mari. 2015. 'The International Relations of the Green Energy Economy in the GCC'. In Jstor. https://www.jstor.org/stable/resrep31026.6.

Mahesh, Sana. 2026. 'Chokepoint Crisis: What the Iran Conflict Means for Global Energy Markets'. International Relations Review, April 28. https://www.irreview.org/articles/2026/4/28/chokepoint-crisis-what-the-iran-conflict-means-for-global-energy-markets.

McKinsey & Company. 2024. 'Global Energy Perspective 2024'. In McKinsey & Company. https://www.mckinsey.com/industries/energy-and-materials/our-insights/global-energy-perspective-2024.

Nawab, Mohaimen. 2025. 'Energy Security and Geopolitics: The Case of the Russia–Ukraine Conflict'. Zenodo 4 (2). https://doi.org/10.5281/zenodo.17583615.

Pant, Harsh V. 2026. 'The US-Iran Crisis and Global Energy Geopolitics'. In Orfonline.Org. OBSERVER RESEARCH FOUNDATION (ORF). https://www.orfonline.org/research/the-us-iran-crisis-and-global-energy-geopolitics.

PwC India, and ASSOCHAM. 2025. 'Augmenting_Role_Gas_Indias_2030_Energy_Mix_Beyond'. September. https://drive.google.com/file/d/13g9L_dTxfypuNfvMNcJgz-j778lvegM_/view.

S&P Global. 2026. 'India Says It Is Diversifying Oil Sources to Bolster Energy Security'. S&P Global, February 16. file:///Users/toshanipadarthi/Downloads/India%20says%20it%20is%20diversifying%20oil%20sources%20to%20bolster%20energy%20security_%20S&P%20Global.pdf.

Sahai, Vrinda. 2026. 'India's Oil Security Strategy: Structural Vulnerabilities and Strategic Choices'. In Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2026/04/indias-oil-security-strategy-structural-vulnerabilities-and-strategic-choices.

San-Akca, Belgin, S. Duygu Sever, and Suhnaz Yilmaz. 2020. 'Does Natural Gas Fuel Civil War? Rethinking Energy Security, International Relations, and Fossil-Fuel Conflict'. Energy Research & Social Science 70 (December). https://doi.org/10.1016/j.erss.2020.101690.

SWP, GRF, and RSIS. 2026. 'How the Iran War Is Reshaping Power and Trade'. In Cfr.Org. Council on Foreign Relations. https://www.cfr.org/articles/how-the-iran-war-is-reshaping-power-and-trade.

U.S. Energy Information Administration. 2026. 'EIA - Annual Energy Outlook 2026'. In Eia.Gov. https://www.eia.gov/outlooks/aeo/.

Yusran. 2025. 'International Relations and Energy Security: Challenges and Opportunities'. Journal of Law, Social Science and Humanities 3 (1): 21–36. https://myjournal.or.id/index.php/JLSSH/article/view/387.

Code-Switching in Shanghai International Schools: Contexts, Motivations, Effects, and Implications By Pete Chen

Abstract

Code-switching, or using many languages to construct the same piece of writing or speech, is a common process for multilingual individuals. Coinciding with the advent of widespread international education is the rise in the population of multilingual international students, particularly in regions like China. This study focuses on studying the causes, circumstances, and consequences of code-switching for multilingual students in Shanghai, China. High schoolers from a Shanghai international school were interviewed, and the data was processed through interpretative phenomenological analysis (IPA) and thematic analysis. Participants discussed topics such as their relative proficiencies in each language which they studied (Chinese and English), their habits of code-switching, and their opinions on the efficacies of code-switching in speech and writing. The results of the analysis showed that code-switching was a process mostly exclusive to informal contexts, and that code-switching commonly occurred when students encountered limitations in their vocabulary. Results also showed that code-switching was more common and preferable in speech than in writing due to a greater prevalence of informal contexts for speaking in student life. Ultimately, the study provided insights into a new and rising culture of international young adults by exploring multilingual students' means of communication in code-switching.

Introduction

Amidst the rise of globalization in the 21st century, an increasing number of primary, middle, and high school students can be found learning globalized curriculums in international schools. Importantly, such students tend to receive multilingual education, given that most intend to pursue higher education abroad (Lee and Wright). This is exemplified in countries like China: as of July 2025, the nation boasts 900+ international schools, with more than 150 of them accepting only foreign passport holders (Hurun Report). In China's international schools, most students study two languages consistently and concurrently—Chinese and English (Poole).

This multilingual education often entails the use of many languages in not just students' communication in the classroom but also their interactions with friends, family, and even society. A common result of this is code-switching, or using many languages in the same piece of writing or speech (Nilep).

Generally, code-switching has been dubbed as a complex but also important process, sometimes creating dilemmas for students (Liang). However, there is currently a limited number of studies exploring the phenomenon of code-switching in East Asia, given that international education is relatively new in the region (Hayhoe et al.). This paper will hence discuss code-switching in Chinese international schools, with a focus on Shanghai in particular—this is since Shanghai leads the nation in number and concentration of international schools (Poole et al.).

Within this geographic scope, the study will explore the following research questions: 1) under what circumstances do international students use code-switching? and 2) what are the effects of code-switching on their communication?

Given the expanding population of multilingual students capable of code-switching, this research is significant for its attempt to better comprehend a new student life and global culture. Furthermore, a more thorough understanding of communication methods among multilingual students can help inform educational improvements or reforms in international schools and institutions.

Methodology

This investigation was conducted with interpretative phenomenology analysis (IPA), a qualitative method that can focus on participants' personal experience with language use (Willis and Harvey). Semi-structured interviews were carried out on 9 multilingual students. Each interview discussed participants' experience in learning and using each language, their specific habits of code-switching, and their opinions on code-switching in both writing and speaking.

Sampling

In this study, participants were limited to one school in Shanghai, China such that differing school systems and student lifestyles would not skew findings on code-switching. Five boys and four girls from grades 10 and 11 (15 to 17 years of age) were interviewed.

Participants were selected such that each had at least 10 years of learning both Chinese and English. This guaranteed considerable proficiency in both languages that allowed substantial code-switching. Also, all selected participants consistently received academic requirements on studying both languages at school. However, participants had varying levels of proficiency in each language, and each may have also preferred a specific language over another.

Data Collection

Interviews were held on an online meeting application (Tencent Meetings), and each session was recorded separately. Then, each recording was processed through a secure AI transcription program, creating precise transcripts. Then, manual correction for grammar, spelling, and possible transcription errors was applied to each transcript.

Data Analysis

The recorded transcripts were analyzed using thematic analysis. Specific codes were extracted from the finalized interview transcripts. Each code summarized a particular point or opinion made by the participant, based on one or more excerpts from the transcript. Responses on different questions or topics made different codes.

Afterwards, codes from different participants were compared and sorted. Codes with similar meaning or content were combined to form sub-themes. These sub-themes were then regrouped under distinct themes, with each theme focusing on a specific topic about participants

(e.g. their degrees of English or Chinese usage or their opinions on code-switching in the classroom). For this investigation, 7 themes were selected for analysis, as shown in *Table 1*:

Themes and Explanations		
No.	**Name of Theme**	**Explanation**
1	Writing: context	The possible circumstances when the participant needs to code-switch in writing
2	Writing: reasons	Why code-switching in writing is necessary and/or preferred in certain cases
3	Writing: effects	Advantages or disadvantages for code-switching in writing
4	Speaking: context	The possible situations when the participant needs to code-switch while speaking
5	Speaking: reasons	Why code-switching in speech is necessary and/or preferred in certain cases
6	Speaking: effects	Advantages or disadvantages for code-switching in speech
7	Writing vs. speaking	Comparing the properties of code-switching in speech and writing

Table 1: A list of the 7 themes and their explanations.

From these 7 themes, a total of 17 sub-themes were analyzed. Data from these sub-themes will be explored.

Results

The data collected in this investigation was divided into two large sections, with each section answering one out of the two proposed research questions: 1) under what circumstances do international students use code-switching? and 2) what are the effects of code-switching on their communication? Specifically, themes #1, #2, #4, and #5 answered the first question, while the remaining themes answered the second.

Answering Question 1: Circumstances for Code-switching

Theme #1 and Sub-themes		
Theme name	**Sub-theme 1**	**Sub-theme 2**
Writing: context	Small need for code-switching in writing for school subjects	Frequently code-switches while texting

Table 2: "Writing: context" and sub-themes.

Most participants stated that code-switching did not occur in academic writing, as shown in sub-theme 1 of *Table 2*. Using many languages in one piece of academic writing is rare, as one participant mentioned:

"*I feel like I only code-switch very casually. In writing, it's usually single-minded."*

However, code-switching in writing may occur when the participants step out of the classroom, in contexts such as communication on social media:

"[When I am] texting someone, I find myself using two languages in the same text."

All in all, participants performed code-switching in writing in mostly only informal contexts.

Theme #2 and Sub-themes		
Theme name	**Sub-theme 1**	**Sub-theme 2**
Writing: reasons	Code-switching occurs due to limited vocabulary, varies with what needs to be expressed	Vocabulary in English supplements vocabulary for Chinese

Table 3: "Writing: reasons" and sub-themes.

After specifying the contexts during which code-switching for writing can occur, participants pointed out that code-switching in writing was usually employed when there was a lack of a better word in a consistent language. This ultimately resulted in a "limited vocabulary", as stated in sub-theme 1 of *Table 3*, and code-switching in writing would be utilized such that this limitation would not hinder expression. One participant stated:

"I think it might be because of my limited vocabulary in some languages, because sometimes when you're texting you find yourself wanting like a word or meaning in one language…so you're switching between languages depending on what you're trying to say."

As implied above, some ideas were best expressed in a particular language. Hence, the purpose of code-switching in writing depended on what was intended to be expressed. For example, if a student wished to quote a Chinese proverb but was initially texting in English, then code-switching from English to Chinese would serve to precisely reflect the meaning of that proverb. Conversely, if that proverb was in English and the writing was initially in Chinese, then the purpose of code-switching in writing would shift.

Notably, the specific purpose of code-switching also depended on the relative language proficiencies of the participant:

"It's hard for me to directly come up with certain words in Chinese, but it's much simpler for me to come up with them in English."

This echoes sub-theme 2 of *Table 3*, and it can be concluded that code-switching in writing occurred due to limited vocabulary, while the specific purpose of such code-switching varied with the content of expression and the student's relative language abilities.

Theme #4 and Sub-themes		
Theme name	**Sub-theme 1**	**Sub-theme 2**
Speaking: context	Uses multiple languages in informal conversations	Code-switching in speech more common for intense conversations

Table 4: "Speaking: context" and sub-themes.

Like in writing, participants also noted that code-switching in speech was limited to informal contexts, such as the "informal conversations" in sub-theme 1 of *Table 4*. It was noted by a participant that code-switching in speech is more common for intense or emphatic (but still informal) conversations:

"I think the most prominent one would be when I'm gossiping with my friends and it's getting heated…[expressing] in many languages makes it better. It's about emphasizing something and getting the best words out."

Because code-switching allowed speakers to choose words across multiple languages, it was mentioned by participants that code-switching in speech could help with more accurate expression (i.e. finding the "best words"). This also helped with delivering points of emphasis in informal conversations. Overall, code-switching in speech occurred also in mostly informal contexts, and it was more common for emphatic conversations.

Theme #5 and Sub-themes			
Theme name	**Sub-theme 1**	**Sub-theme 2**	**Sub-theme 3**
Speaking: reasons	Code-switches in speech because recipients can still comprehend	Code-switching in speech comes from experience in multilingual communities	Code-switching in speech also depends on speech habits

Table 5: "Speaking: reasons" and sub-themes.

Multiple reasons were accounted for code-switching, as shown in *Table 5*. Many participants mentioned that they performed code-switching in speech because they knew that their recipients, or those whom they converse with, would be able to understand them:

"There's just no pressure to stick to one language. So when you switch from Chinese to English, it doesn't matter because the same people will understand you."

Similarly, code-switching in speech also occurred due to exposure to multilingual communities. Most participants explained that their school's bilingual environment was a source, where classes were sometimes taught in both English and Chinese:

"The teachers are instructed to teach in English…if they aren't the best at English, then they kind of just don't care and speak in Chinese."

Some participants also experienced code-switching from living in their countries of nationality, such as Singapore:

"[Code-switching is] part of Singaporean speaking. If you go to Singapore, everyone speaks English with their mother tongue."

In these cases, it was not simply a lack of "pressure" that allowed code-switching in speech, but a normalization of code-switching that encouraged it.

Furthermore, some participants also pointed out that code-switching in speech can be habitual. As stated by one participant, some expressions were simply more interesting to deliver in a certain language:

"Maybe there's a certain phrase or word pattern that is in English and it's funnier to say it that way, so even if I'm talking in Chinese, I would want to [switch languages]."

Hence, code-switching in speech may occur because of mutual understanding among speakers and experience in a multilingual community, but it may also happen due to certain speech habits.

Answering Question 2: Effects Of Code-Switching

Theme #3 and Sub-themes			
Theme name	**Sub-theme 1**	**Sub-theme 2**	**Sub-theme 3**
Writing: effects	Grammar and word use incompatible for code-switching in writing	Limited vocabulary still hinders clarity, even with code-switching	Code-switching in writing helps with better language construction

Table 6: "Writing: effects" and sub-themes.

A common consensus reached among many participants' answers was that code-switching in writing would almost inevitably result in a grammatical mismatch. In fact, some participants reported that they were sometimes discouraged from code-switching in writing because of this mismatch:

"If you're using English to write Chinese, it's just not the same. They have different grammatical structures. They have different usage of words. It doesn't really work."

Ironically, another drawback to code-switching originated from limited vocabulary, as mentioned in sub-theme 2 of *Table 6*. As stated by a participant, there were times when switching to a different language didn't necessarily result in clearer expression:

"I don't have enough vocabulary, or I just can't find a way to express something beyond a very basic term."

Thus, code-switching had drawbacks of not only grammatical incompatibility but also a potential loss of clarity in communication. However, it was agreed by most participants that code-switching in writing could still be beneficial to language construction.

"I think [code-switching in writing] does help me express myself better… [the two languages] have a kind of 'filling in' for each other, so there's always something that I can use, rather than having to be limited to what I have for one language."

In many cases, switching to a certain language can supplement a lack of vocabulary in another, resulting in more complete meaning, given that recipients can also comprehend both languages.

Theme #6 and Sub-themes			
Theme name	**Sub-theme 1**	**Sub-theme 2**	**Sub-theme 3**
Speaking: effects	Code-switching in speech implies poor language ability	Code-switching is more helpful for less proficient language	In code-switching for speech, two languages can make up for each other's "inefficiencies"

Table 7: "Speaking: effects" and sub-themes.

When it came to code-switching in speech, participants also pointed out many positives and negatives. Like in sub-theme 1 of *Table 7*, participants stated that code-switching in speech often exposed grammatical incompatibilities between Chinese and English, which would sometimes underrepresent speakers' linguistic abilities, implying that the student was not capable of communicating in a particular language:

"[Code-switching] makes you seem like you're worse at Chinese than you actually are, in my opinion."

Participants also noted that code-switching in speech tended to be more noticeably beneficial for their less proficient language. For example, according to a participant who reported being less proficient in Chinese:

"[Code-switching is helpful] for one of the times when I want to use Chinese, when I can't come up with a word to say."

Since code-switching in speech often arose from limited vocabulary, it is thus also reasonable that this action became more helpful with a more limited vocabulary.

Importantly, code-switching in speech was found to not only help with word choice issues but also save time and improve efficiency. An epitome of this difference in efficiency across languages would be the English and Chinese words "P.E." and "体育课"—both refer to "physical education class", but the former version in English is more concise (with less syllables) and more memorable in communication. As stated by one participant, code-switching in speech yields the more efficient way of phrasing one's message:

"*I think to a certain extent…there are words that Chinese encapsulates much more simply than English and also words that English expresses much more cleanly than Chinese."*

In this sense, code-switching also benefited students by offering faster and clearer alternatives in speech.

Theme #7 and Sub-themes		
Theme name	**Sub-theme 1**	**Sub-theme 2**
Writing vs. speaking	Code-switching more prevalent in speech than writing	Code-switching in speech more comfortable/smooth because of less formality required

Table 8: "Writing vs. Speaking" and sub-themes.

As concluded by most participants, code-switching was more common in speech than in writing:

"I code-switch a lot more when I talk than when I write. A lot more."

This was because of a greater number of informal contexts for speech than for writing, and hence more contexts were appropriate for code-switching. This also relates to sub-theme 2 of *Table 8*, in which participants believed that code-switching in speech was a more comfortable process due to less pressure for formality:

"I think there's less difficulties when it's speaking because it's less formal and you don't have stress or requirements."

Since code-switching was more prominent in speaking than in writing, it can be concluded that the effects of the former tend to be more widespread than those of the latter.

Discussion

This study answered two research questions: 1) under what circumstances do international students use code-switching? and 2) what are the effects of code-switching on their communication? Specifically, it was found that code-switching was performed mainly in informal contexts and due to limited vocabulary in each language. Consequently, code-switching often resulted in clearer and more efficient expression. Because of a greater number of informal contexts for speaking, code-switching in speech was more common than in writing.

The findings of this study were consistent with previous research. Zia et al. found that code-switching among Pakistani students who were bilingual in Urdu and English was strongly influenced by students' relative linguistic proficiencies and their school environments. Moreover, Li Liu found that, for many international students switching between Spanish and English, code-switching allowed clearer and more natural expression of emotions.

The consistent findings from these studies imply that the benefits of code-switching apply not only to a Chinese-English bilingual student environment, but also to international students in other contexts and backgrounds.

The significance of this study lies in its implications for language and linguistic education. Code-switchers exist partially because their vocabulary in a single language is not enough to satisfy their daily needs, but also because each language contains phrases and words that they find too "inefficient" and slow to use. This suggests that future education should include teaching multilingual students how to express themselves efficiently.

For international schools, like the one attended by the participants in this study, a better understanding of code-switching in the student community can inform teaching strategies. For parents, a thorough understanding of code-switching can help foster parent-child bonds and better communication at home.

Conclusion

This study explored the experience of code-switching among international students. Future studies can focus on the impact of teachers' behavior on students' code-switching strategies. Studies may also investigate and compare possible policies (e.g. code-switching tolerance and/or restrictions) in the school environment to facilitate better communication for multilingual students. Ultimately, what a young code-switcher wants is smoother talks and more understanding, and it is research that can inform how this can be better achieved.

Works Cited

Hayhoe, Ruth, et al. Educational History, Models, and Initiatives. 1st ed., Berkshire, 2013, https://www.jstor.org/stable/j.ctt1j0pszw. JSTOR. Accessed 17 Aug. 2026.

"Hurun Report - Info - Hurun Education Top International Schools in China 2025." Hurun.Net, 2025, https://www.hurun.net/en-US/Info/Detail?num=G822XEBEK5DL&utm_. Accessed 16 Aug. 2026.

Liang, Xiaoping. "Identity and language functions: High school Chinese immigrant students' code-switching dilemmas in ESL classes." Journal of Language, Identity, and Education 5.2 (2006): 143-167.

Lee, Moosung, and Ewan Wright. "Moving from Elite International Schools to the World's Elite Universities." International Journal of Comparative Education and Development, vol. 18, no. 2, May 2016, pp. 120–36 https://doi.org/10.1108/ijced-01-2016-0002.

LI LIU, Edson. Code-Switching and Bilingual Identity: How Bilingual High School Students Use Language to Express Emotion and Cultural Identity. Elsevier BV, 2026, https://doi.org/10.2139/ssrn.6375598.

Nilep, Chad. "'Code Switching' in Sociocultural Linguistics." Colorado Research in Linguistics, vol. 19, no. 19, 2006, https://doi.org/10.25810/hnq4-jv62. Accessed 16 Aug. 2026.

Poole, Adam. "Decoupling Chinese Internationalised Schools from Normative Constructions of the International School." Compare: A Journal of Comparative and International Education, vol. 50, no. 3, Oct. 2020, pp. 447–54, https://doi.org/10.1080/03057925.2019.1682839.

Poole, Adam, et al. "'We're Away From Everything': Understanding the Struggles Faced by Internationalized Schools in Non-Urban Contexts in China." SAGE Open, vol. 12, no. 1, Jan. 2022, p. 215824402210810, https://doi.org/10.1177/21582440221081026.

Willis, Rohan K., and Sharon Harvey. "Doing Interpretative Phenomenological Analysis (IPA) in Applied Linguistics Research: A Field-Specific Guide." Research Methods in Applied Linguistics, vol. 4, no. 3, Elsevier, Nov. 2025, p. 100279, https://doi.org/10.1016/j.rmal.2025.100279. Accessed 16 Aug. 2026.

Zia, T., S. Batool, and N. A. Qureshi. "Analyzing the sociolinguistic factors influencing code-switching among students in bilingual classrooms." International Journal of Contemporary Issues in Social Sciences 3.3 (2024): 3212-3231

The Bilingual Brain: Does Language Dominance Affect Cognitive Flexibility?
By Ziyu Wang

Introduction

It is well researched that bilingual experience is associated with better performance on particular measures of executive control. Although later studies show this advantage is context-dependent and not universally observed (Lehtonen et al. 396). The term "bilingual advantage" is often used to describe proposed cognitive or neural benefits associated with bilingual experience. Some studies have reported a later age of dementia symptom presentation among bilinguals, but this association remains debated and should not be interpreted as evidence that bilingualism universally prevents dementia (Olsen et al. 131). Diffusion tensor imaging (DTI) has been used to examine associations between bilingual experience and white-matter microstructure in older adults (Anderson et al. 143). Bilinguals showed greater white matter integrity (seen through higher fractional anisotropy and lower mean diffusivity) compared to monolinguals which allows a more efficient subnetwork for transferring information between language and control centres. Moreover, some neuroimaging studies claim bilinguals show greater grey matter volume (possibly due to less developmental pruning) compared to monolinguals. The regions most pronounced were those left-lateralised regions crucial for language processing, cognitive shifting and executive control: Inferior Frontal Gyrus, Superior Frontal Cortex, Inferior and Superior Parietal Cortices and Precuneus (Pliatsikas et al. 2131). This is hypothesised to function as a 'cognitive reserve,' potentially mitigating age-related neurodegeneration (Anderson et al. 150). Although the differences in language dominance in bilinguals raise questions about whether they affect executive function in any way (Treffers-Daller 375). This paper aims to explore how language dominance in bilinguals affects neurocognition.

Language dominance in bilingual profiles

In Treffers-Daller's framework, bilingualism is clearly treated as a spectrum, but this does not deny that some individuals can legitimately be described as balanced or dominant bilinguals. The key point is that dominance is defined relatively, along two core dimensions: language proficiency ("what a person can do" in each language) and language use ("what a person typically does" in each language), and these do not always line up perfectly. Most bilinguals show a stronger and a weaker language somewhere in this multidimensional space, but some fall sufficiently close to parity on chosen measures to be categorised as balanced (Treffers-Daller 376).

Balanced bilinguals, in this paper, are those whose two languages are so similar in strength that they do not exceed an agreed dominance threshold on the measures used. For example, Genesee et al. measured language proficiency with indices such as mean length of utterance, lexical variety and proportion of multimorphemic utterances (Genesee et al. 616). This leads to the belief that balanced bilinguals did not clearly favour one language, dominant

bilinguals had a stronger diversion. Bedore et al.'s research had very large samples and computed difference scores for semantic and morphosyntactic performance; children whose scores in the two languages were very close could be put into a "balanced" category, in contrast to those who were clearly Spanish- or English-dominant (Bedore et al. 616–629). Unsworth similarly defined dominance via differences larger than about one standard deviation on linguistic measures; children who did not cross that difference threshold were treated as more balanced, whereas those who had a clear difference in the manipulation of the two languages were classified as dominant bilinguals (Unsworth 157).

Treffers-Daller stresses that there is no single universal cutoff, so the exact numerical boundary between "balanced" and "dominant" is a methodological decision that must be made explicit. Moreover, balance can be global (no clear dominance across the chosen proficiency and use measures) or domain-specific (a person might be dominant in one language for school/work but effectively balanced, or even dominant in the other language, for home or emotional topics). Because bilingual knowledge is "distributed" across domains, a person may be balanced on some skills, in some contexts, and not in others. Ultimately, this means balanced bilinguals are not an impossibility; rather, they are defined as individuals whose relative proficiency and use in both languages are close enough—based on clear, study-specific criteria—that neither system can be considered meaningfully stronger overall, even though perfect equality across all skills, domains, and situations remains rare (Treffers-Daller 377).

Development of the bilingual profiles

Children who grow up in multilingual households are frequently on a trajectory toward bilingualism because they are exposed to two languages from the earliest stages of development in everyday, meaningful interactions. Early in life, however, their abilities in these two languages are rarely perfectly balanced. Typically, one language is used more frequently—home, caregivers, siblings, extended family—and is therefore practised more intensively. Within an incremental learning framework, where language is treated less as a fixed set of rules and more as a skill that is strengthened through repeated use, this early asymmetry in practice naturally yields an initial dominance in the more frequently used language (Oppenheim et al. 21-22). At this stage, children can be described as dominant bilinguals, not because the other language is absent, but because their abilities may differ across the two languages depending on language use and the measures used to assess proficiency. As the child's environment changes, however, this profile can shift quite naturally (Bedore et al. 621-623; Oppenheim et al. 21). When both languages are pulled into a wider range of daily contexts—like home, school, peers, and community life—each system begins receiving its own steady stream of practice. Incremental learning accounts suggest that every instance of successful word retrieval or comprehension in either language slightly strengthens the underlying form–meaning connections. This slow accumulation helps stabilise their relative strengths, allowing a more balanced bilingual profile to take shape as daily usage becomes more evenly distributed. Longitudinal evidence with Spanish–English children transitioning from mostly Spanish-speaking homes into English-dominant schools illustrates this

process: as the children's English use increased, their English vocabulary size and speed of word retrieval grew to the point that English became their more proficient language around third grade. Yet their Spanish skills also continued to improve rather than collapse, demonstrating that increased practice in one language does not automatically cause loss of the other; instead, the eventual balance or dominance reflects the cumulative history of practice in each system (Oppenheim et al. 41).

A different developmental pathway arises when environmental exposure remains strongly imbalanced across childhood. If a child from a bilingual home moves into a school system and peer network that overwhelmingly uses one societal language, daily opportunities to use the L1 may reduce substantially. From an incremental learning perspective, this widening gap means that mappings in the majority language are continually reinforced through dense, high-frequency use, while connections in the minority or home language are left largely unpractised. Empirically, this dynamic can lead to cases where the second, societally dominant language not only catches up to the first language in vocabulary and retrieval speed but eventually surpasses it entirely, becoming the child's dominant tongue. Meanwhile, the first language might either continue to grow slowly—assuming it is still used at home and in some community circles—or it may show clear signs of stagnation or attrition if practice becomes genuinely rare. Because the very same mechanisms that drive learning also support "competitive unlearning," each time a concept is expressed in the dominant language, mappings in that code are strengthened at the expense of the weaker, less-frequently accessed pathways in the other language. Over long stretches of time, this constant competition can reduce the accessibility of words in the less-practised language, resulting in slower retrieval and apparent "forgetting," particularly for items that are not regularly reinforced. The longitudinal data from Spanish–English children show both possibilities: many children increased their English use enough that English became clearly dominant, yet their Spanish performance still improved when there was sufficient continuing practice, indicating that attrition is not inevitable but depends on how much the weaker language continues to be exercised (Oppenheim et al. 20, 40-42).

A parallel set of dynamics can be seen when individuals who start out as monolinguals begin to learn a new language later in life. Initially, the native language has a dense, well-practised network of mappings that support rapid and reliable retrieval, whereas the newly learned language begins as a sparse, fragile system in which words are slower and harder to access. From an incremental learning perspective, this initial state corresponds to a strongly dominant L1 and a weak L2. If the learner then enters an environment that demands frequent, meaningful use of the L2—for instance, by immigrating to a country where the L2 is the majority language—the balance of practice can shift dramatically. Under these immersive conditions, the learner is constantly forced to retrieve L2 words quickly and accurately during real-time interactions, which rapidly accelerates the strengthening of L2 mappings. Over time, this expanding L2 vocabulary and faster processing can bring L2 performance close to, or even above, L1 benchmarks on certain tasks, shifting dominance or producing a more balanced

profile—much like the Spanish–English children whose English proficiency overtook their Spanish while their native tongue still improved with ongoing use (Oppenheim et al. 42).

In contrast, if L2 exposure remains limited to a few classroom hours per week, as is common in foreign language instruction, the opportunities for incremental strengthening of L2 mappings are sparse and relatively decontextualised. The learner's native language continues to be used for most daily communication, education, and internal thought, so its mappings stay highly practised and efficient. In this situation, the L2 may remain comparatively less accessible: vocabulary growth may be slower, and word retrieval may remain more effortful. Within an incremental-learning account, limited and low-frequency L2 use may be insufficient to reduce the advantage of the more frequently practised language; however, outcomes vary according to the learner's exposure, age of acquisition, and opportunities for meaningful L2 use (Oppenheim et al. 40-43).

Looking across all of these trajectories—whether simultaneous bilinguals in multilingual homes, children moving into majority-language schools, or monolinguals acquiring an L2 later in life—the same core principles hold true. Bilingual abilities at any given moment simply reflect the cumulative amount and real-world distribution of practice in each language. Both gains and losses in proficiency stem from the exact same experience-driven learning mechanisms that constantly fine-tune the strength of competing linguistic representations over time (Oppenheim et al. 40-42).

Neural differences amongst bilingual profiles

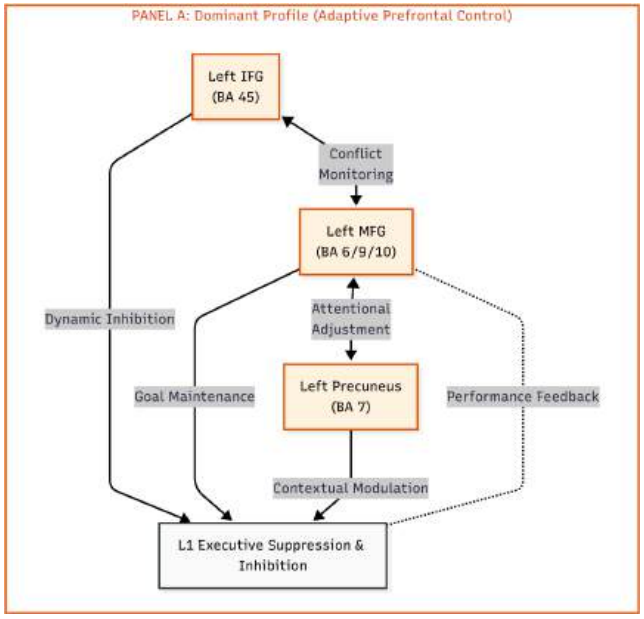
Extensive work on bilingualism-induced neuroplasticity shows that the brains of bilinguals and monolinguals differ not only in traditional "language areas" but across widespread networks that support learning, control, and communication between regions. Magnetic resonance imaging studies indicate that acquiring and using an additional language produces structural changes in grey matter (in cortical, subcortical, and cerebellar areas) as well as white-matter pathways linking these regions—and these modifications evolve dynamically rather than representing a fixed endpoint (Pliatsikas 459-460). Early in the language-learning process, there is a visible expansion of cortical regions involved in vocabulary, phonological, and semantic processing, alongside executive control. This is particularly noticeable in parietal and temporal areas, such as the inferior and superior parietal lobules, anterior temporal regions, and auditory cortices, alongside frontal areas like the inferior and middle frontal gyri and the anterior cingulate cortex. This structural pattern mirrors the heavy cognitive demands of rapidly acquiring new lexical items and learning to select between competing representations (Pliatsikas 464).

With continued exposure, and especially through immersion, these initial cortical enlargements tend to normalise, while subcortical structures tied to language production and control—such as the caudate nucleus, putamen, globus pallidus and thalamus—undergo structural volume adaptations (Korenar et al. 2). At the same time, white-matter tracts connecting frontal, temporal, and parietal regions (including both ventral pathways for semantic processing

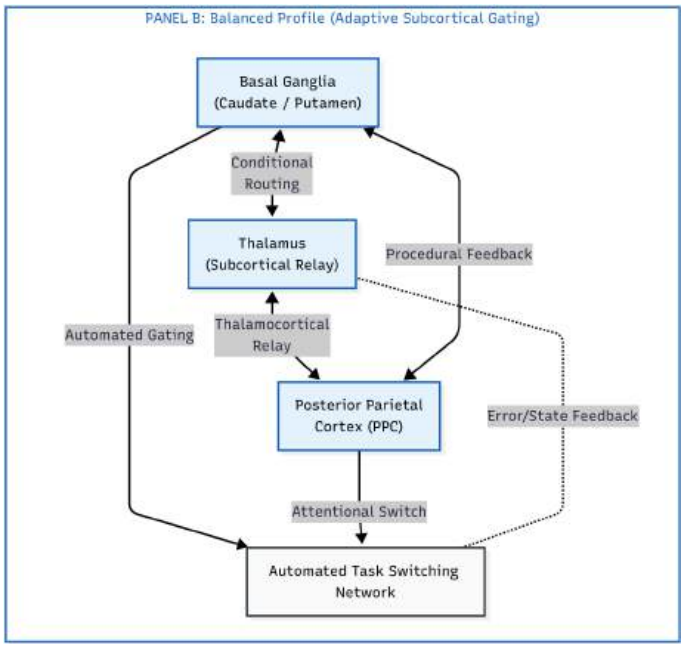
and dorsal pathways for phonology and complex syntax), alongside interhemispheric tracts like the corpus callosum, show decreased diffusivity, indicating more efficient structural connectivity. This transition from widespread cortical engagement to strengthened subcortical circuits and optimised connectivity is widely taken to index the growing efficiency and automation of language control. Because bilinguals must continuously learn new words, inhibit non-target lexical, phonological, and motor programmes, and switch between languages as their environment demands, the brain reorganises on the fly, allowing these intense control operations to run with less costly, more specialised, and better coordinated neural resources (Pliatsikas 464-465).

Although broad patterns of bilingual brain organisation have been widely documented, much less is known about how the brain differs between balanced bilinguals, who use their languages with relatively equal proficiency, and dominant or unbalanced bilinguals, whose abilities are clearly stronger in one language than the other. To address this gap, Wang's activation likelihood estimation (ALE) meta-analysis pooled functional MRI studies that specifically examined bilingual language control and separated participants by language balance (Wang et al.).

As illustrated in Figure 1, the dominant bilingual profile relies on a high-demand prefrontal control network (Panel A), whereas the balanced profile transitions language selection to an automated, cortico-striatal gating circuit (Panel B).



Panel A: Dominant Profile (Adaptive Prefrontal Control) Schematic illustration of top-down executive control networks engaged during asymmetrical language suppression. Cortical nodes—including the Left Inferior Frontal Gyrus (IFG; BA 45), Left Middle Frontal Gyrus (MFG; BA 6/9/10), and Left Precuneus (BA 7)—operate through reciprocal cross-talk (*conflict monitoring* and *attentional adjustment*) to drive active L1 executive suppression and inhibition. Dynamic inhibition and goal-maintenance pathways continuously adjust prefrontal output, regulated by feedback loops that evaluate task performance in real time.



Panel B: Balanced Profile (Adaptive Subcortical Gating) Schematic illustration of proceduralised subcortical-parietal gating loops engaged during balanced dual-language switching. Recurrent cortico-striatal circuits involving the Basal Ganglia (Caudate/Putamen), Thalamus, and Posterior Parietal Cortex (PPC) execute automated task switching with minimal prefrontal demand. The network relies on conditional routing and thalamocortical relays, paired with procedural and error/state feedback loops that dynamically adjust gating thresholds during active language selection.

By statistically combining activation peaks across many studies, the analysis identified both common and distinct neural patterns for these two bilingual profiles. A conjunction analysis first showed a shared, domain-general control network, with overlapping activation in the left precentral gyrus and left medial frontal gyrus in both balanced and unbalanced bilinguals during language control tasks, indicating that both groups recruit a common core system to manage competition between their languages. However, when the two groups were directly contrasted, a clear asymmetry emerged. Balanced bilinguals showed no regions of significantly greater activation convergence than unbalanced bilinguals, suggesting that balanced bilingualism does not simply add extra control demands at the neural level (Wang et al.)

Instead, unbalanced bilinguals exhibited significantly stronger convergence of activation in the left middle frontal gyrus (Brodmann areas 6, 9, 10), the left inferior frontal gyrus (BA 45), and the left precuneus (BA 7) during tasks like language switching, which place high demands on selecting and shifting between codes (Wang et al.). These regions are typically associated with higher-order control, selection, and internally directed processing, and their stronger engagement in unbalanced bilinguals suggests that unequal proficiency makes coordinating two languages more effortful. Because dominant bilinguals must exert greater control to manage interference from their weaker language and to switch efficiently, they appear to draw more heavily on domain-general cognitive control resources than balanced bilinguals, whose more evenly

developed language systems can be regulated with less reliance on these additional frontal and parietal mechanisms (Wang et al.).

Synthesising the literature reveals a fundamental divergence between behavioural and neurofunctional meta-analyses. Broad behavioural meta-analyses frequently report negligible or near-zero overall effect sizes for domain-general cognitive advantages, often attributing observed group differences to publication bias or socio-demographic confounds. Conversely, neuroimaging meta-analyses demonstrate robust spatial convergence in fronto-parietal and subcortical activation across tasks. Resolving this discrepancy requires recognising that aggregate behavioural meta-analyses often obscure domain-specific effects by grouping heterogeneous executive tasks, whereas neural data consistently reflect altered cognitive processing strategies even when behavioural output appears equivalent.

Adaptive Control Hypothesis (ACH)

Green and Abutalebi's adaptive control hypothesis starts from the idea that both speech comprehension and production are governed by general control processes, and that in bilinguals these language control processes themselves are not fixed but adapt to the recurrent demands of everyday language use (Green & Abutalebi 515; Lai & O'Brien). "Adaptation" here means that repeated experience in a given pattern of language use changes parameters of the control system: its neural capacity or efficiency, and the way distinct control processes are coordinated or cascaded together within specific neural circuits (Green & Abutalebi 515). Thus, the core claim is stronger than saying that bilinguals merely "use more control": over time, the control system is reshaped by the interactional context in which the two languages are used (Green & Abutalebi 515; Ng & Yang 521).

The authors distinguish eight key control processes involved in bilingual language production: goal maintenance (keeping the intended language and message in mind), conflict monitoring (detecting competition between languages), interference suppression (reducing activation of the non-target language), salient cue detection (picking up cues that signal which language is appropriate), selective response inhibition (stopping a prepared response in the wrong language), task disengagement and task engagement (switching from one language schema to another), and opportunistic planning (preparing future utterances when circumstances allow) (Green & Abutalebi 515; Calabria et al. 222).

They then analyse how three real-world interactional contexts place different demands on this set: a single-language context (only one language used in each setting), a dual-language context (both languages used in the same setting with different interlocutors), and a dense code-switching context (frequent, fluid mixing within the same conversation) (Green & Abutalebi 517-519; Lai & O'Brien). In single-language contexts, activating a language schema and inhibiting the other mainly recruits goal maintenance, conflict monitoring, and interference suppression. Dual-language contexts additionally require frequent task disengagement/engagement and selective response inhibition as speakers switch between languages for different interlocutors (Green & Abutalebi 518). Dense code-switching, by

contrast, fosters cooperative use of both languages and relies more on salient cue detection and opportunistic planning than on strict inhibition (Green & Abutalebi 519).

On this basis, the hypothesis predicts that recurrent use of a given interactional context will lead to specific adaptive changes in the neural regions and circuits that implement these processes (Green & Abutalebi 520; Calabria et al. 228). For example, sustained experience in a dual-language context is expected to tune a circuit that links conflict monitoring, interference suppression, selective inhibition, and task switching so as to circumvent a "control dilemma" where excessive inhibition would slow communication, but insufficient inhibition would allow interference (Green & Abutalebi 522-523). These adaptations should be observable as changes in activity and connectivity within a distributed language control network, and, under some conditions, may generalise to nonverbal executive control tasks (Green & Abutalebi 525). Crucially, the authors argue that strong tests of the hypothesis require assessing multiple control processes, in both language and non-language tasks, within the same individuals, and relating behavioural patterns and neuroimaging evidence directly to detailed characterisations of each bilingual's typical interactional contexts (Green & Abutalebi 526; Ng & Yang 522).

Conditional Routing Model (CRM)

The conditional routing model accounts for general cognitive coordination by highlighting the basal ganglia's role as a flexible gating system (Stocco et al. 541; Yamasaki et al. 1). This system dynamically directs neural information flow to the prefrontal cortex based on current linguistic demands, enabling bilinguals to selectively activate the appropriate language network in response to contextual cues and efficiently manage cross-linguistic competition (Stocco et al. 543-545).

When theoretically extended to dominant versus balanced bilinguals, this model helps explain distinct routing demands and differences in neural efficiency. Dominant bilinguals—whose proficiency is heavily skewed toward one language—tend to rely much more on conditional gating to suppress interference from their stronger language when attempting to use the weaker one (Calabria et al. 223-224; Yamasaki et al. 2). Because the dominant language features deeply entrenched and automatically accessible representations, the basal ganglia's gating mechanism must exert greater control over routing relevant information to the prefrontal cortex to block dominant-language intrusions (Stocco et al. 546; Calabria et al. 224). This aligns with activation studies showing that dominant bilinguals exhibit heightened engagement in both the basal ganglia and the anterior cingulate cortex during language switching and preparation phases, reflecting the increased gating and monitoring demands required to override native-language interference (Calabria et al. 224-225).

Balanced bilinguals, on the other hand, show more efficient gating mechanisms due to their evenly weighted language representations and the lack of a pronounced dominance effect (Yamasaki et al. 2, 7-8). This parity results in smoother conditional routing with less need for aggressive suppression of either language. Consequently, balanced bilinguals frequently display less extensive or less effortful recruitment of basal ganglia gating circuits during language control

tasks, managing competition through more subtle, automatic gating adjustments (Yamasaki et al. 7-9).

Extrapolating from this model suggests that these differing gating dynamics stem directly from how bilingual experience shapes striatal-prefrontal connectivity over time (Stocco et al. 550). Prolonged, balanced bilingual activation trains the basal ganglia gating system to optimise routing, which ultimately reduces the cognitive cost and neural effort needed for language control (Yamasaki et al. 8-9). For dominant bilinguals, however, the gating system stays actively engaged to compensate for unbalanced language strength, keeping a higher demand on cognitive and neural resources within the control networks (Calabria et al. 224; Yamasaki et al. 8).

Effective coordination between the basal ganglia, prefrontal cortex, and anterior cingulate cortex also supports the attentional control necessary during language switching or preparation. These brain regions work collectively to monitor conflict, implement response selection, and regulate the conditional updating of task-relevant language rules (Stocco et al. 544-545). This is why dominant bilinguals often show greater anterior cingulate activation—reflecting higher conflict-monitoring demands when preparing to use their weaker code—while balanced bilinguals benefit from stable gating and less frequent conflict detection (Calabria et al. 224-225; Yamasaki et al. 8).

Ultimately, the conditional routing model offers a clear, neurocomputational explanation for how dominant and balanced bilinguals differ in language control neural dynamics (Stocco et al. 541). It captures how individual differences in language proficiency balance modulate the degree and efficiency of gating between cortical regions, directly influencing the cognitive and neural costs of managing bilingual language competition (Yamasaki et al. 8-9).

An unaddressed theoretical paradox regarding interactional contexts and control demands rises. While dual-language environments are theorised to demand heightened interference suppression, dense code-switching environments are proposed to rely on opportunistic planning without strict inhibition. However, cross-study comparisons reveal inconsistent effect sizes across task-switching paradigms, challenging the assumption that specific environmental contexts yield predictable executive adaptations. The paradox lies in whether code-switching enhances selective inhibition through continuous language management or diminishes it through cooperative activation, a conceptual contradiction that current empirical literature has yet to resolve systematically. This contradiction highlights a critical theoretical gap.

Empirical literature supporting basal ganglia gating mechanisms reveals significant variability in reported effect sizes across cognitive control paradigms. While neurocomputational models presume that gating efficiency scales closely with language balance, behavioral and neuroimaging studies show fluctuating effect magnitudes when comparing dominant and balanced bilinguals. These variations demand further study. Empirical data remains notably mixed. This exposes a key theoretical gap: the model does not yet account for how individual differences in baseline executive capacity modulate subcortical routing efficiency, leaving open

the question of whether observed neural differences reflect true structural adaptations or pre-existing cognitive divergences.

Bilingual Anterior-to-Posterior and Subcortical Shift (BAPSS) Hypothesis

The neural correlates of cognitive processing in monolinguals and bilinguals reveal distinct patterns of brain activation, reflecting differences in how language experience shapes cognitive control networks. Grundy, Anderson, and Bialystok provide a comprehensive integrative review that highlights a key neural adaptation in bilingual individuals: a shift from anterior to posterior and subcortical brain regions during cognitive tasks. This shift is critical to understanding the bilingual brain's efficiency and flexibility in managing multiple languages (Grundy et al. 184).

In language acquisition, cognitive processing primarily engages the frontal cortex, especially the prefrontal cortex, which is associated with executive functions such as attention, inhibition, and working memory (Grundy et al. 185; DeLuca et al. 4-6). These anterior regions are responsible for top-down control mechanisms that regulate goal-directed behaviour. In contrast, bilinguals demonstrate a more distributed neural network involving posterior cortical areas and subcortical structures, such as the basal ganglia, which play a vital role in procedural learning and automatic processing (Grundy et al. 184-185; Yee et al. 716).

This anterior-to-posterior and subcortical shift suggests that bilingual experience leads to a fundamental reorganisation of cognitive control systems, moving toward a heavier reliance on subcortical pathways and posterior cortical regions to support more efficient, automatic control. For example, the basal ganglia contribute to language switching and conflict resolution by facilitating rapid selection and inhibition of competing linguistic representations. This subcortical involvement directly reduces the cognitive load placed on the prefrontal cortex during demanding tasks (Grundy et al. 188-189).

Neuroimaging studies reviewed by Grundy et al. demonstrate that bilinguals often show reduced activation in frontal control regions during executive function tasks compared to monolinguals, alongside increased engagement of posterior parietal areas. This pattern indicates that bilinguals may utilise more efficient neural strategies, relying on brain regions associated with attentional orienting and sensorimotor integration. The posterior parietal cortex, implicated in attentional control and spatial processing, may enhance bilinguals' ability to manage competing stimuli and switch between tasks more fluidly (Grundy et al. 189-191).

Furthermore, electrophysiological evidence supports this neural shift, with bilinguals exhibiting different event-related potential (ERP) components indicative of altered timing and allocation of cognitive resources. These findings align with behavioural data showing bilingual advantages in tasks requiring conflict monitoring and cognitive flexibility (Grundy et al. 191-193).

The way the brain manages multiple languages evolves significantly as a speaker progresses from being dominant in one language to achieving a more balanced bilingual profile. For dominant bilinguals—those who are still in the early stages of learning a second language or

who use one language much more frequently than the other—processing is primarily managed by anterior brain structures. Because the second language is not yet fully automated, the frontal regions, such as the prefrontal cortex, must provide intense top-down executive control to navigate the effortful demands of speech and to suppress interference from the more familiar language (Grundy et al. 194; DeLuca et al. 3-5). In this profile, neural activity is concentrated in areas dedicated to active monitoring, inhibition, and suppression of the L1, making language tasks cognitively demanding and metabolically expensive (Grundy et al. 195).

As individuals transition into balanced bilingualism through years of consistent and frequent use of both languages, the brain undergoes a functional reorganisation known as the BAPSS framework. This change moves the primary neural workload away from frontal control centres and directs it toward deeper subcortical structures, like the basal ganglia and thalamus, as well as posterior cortical areas. For these balanced speakers, language management becomes a more streamlined, bottom-up process that runs on procedural and automatic pathways rather than constant executive oversight. This structural adaptation shows up in the complexity of neural signalling; balanced bilinguals demonstrate more intricate brain signals in the occipital regions compared to monolinguals, indicating that their sensory and perceptual networks have been structurally refined by their ongoing linguistic experience (Grundy et al. 195-196).

This reorganisation fundamentally changes the timing and efficiency of cognitive processing. Balanced bilinguals tend to detect and resolve linguistic conflicts much earlier in the processing stream, allowing for more rapid and confident classification of linguistic information. Because the brain resolves these conflicts during early processing stages, it requires fewer resources for later error-monitoring, which preserves overall cognitive capacity for the L2, facilitating more automated, proceduralised language execution (Grundy et al. 197).

Ultimately, this integrative review emphasises that the neural shift is not a static byproduct of knowing two languages, but a reflection of dynamic, experience-dependent plasticity within cognitive control networks. Moving from anterior frontal regions to posterior and subcortical areas optimises neural efficiency, reduces the reliance on metabolically expensive frontal structures, and builds automaticity (Grundy et al. 197-198). In dominant bilinguals, management relies heavily on active inhibition and top-down executive control, while balanced bilinguals transition toward automated, habitual language processing (DeLuca et al. 4-5). Mapping these neural correlates provides a clearer mechanistic look at how bilingual experience shapes brain function, while offering broader insights into cognitive ageing and neuroplasticity (Grundy et al. 198).

Reported effect sizes for language balance and executive control vary considerably across behavioural task-switching paradigms. Standardised effect magnitudes for switch costs are consistently larger in paradigms requiring active schema reconfiguration than in simple conflict-monitoring tasks (Grundy et al., 183–201). However, prior literature rarely isolates the specific effect size of language balance from confounding background variables such as age of acquisition or socio-economic status. Because studies fluctuate between small ($d \approx 0.15$) and

moderate ($d \approx 0.45$) effect sizes across different tasks, the true magnitude of balance-induced cognitive plasticity remains inadequately quantified.

Dynamic Restructuring Model (DRM)

The Dynamic Restructuring Model (DRM) offers a comprehensive framework for understanding the complex, variable structural plasticity observed in the bilingual brain. Moving away from earlier views that treated bilingual adaptations as static or uniform, the DRM posits that structural changes are dynamic, temporally staged, and highly dependent on the nature and extent of language experiences over time (Pliatsikas 464–465). This model helps reconcile seemingly inconsistent findings in current literature, such as conflicting reports of both increases and decreases in grey and white matter volumes across different regions (Pliatsikas 459).

At the core of the DRM is the principle that neuroplastic changes related to bilingualism unfold through multiple stages that reflect the evolving demands of language learning and use. Initially, as a person begins to acquire and actively use a second language, grey matter volume expands in regions critical for language processing and cognitive control. This initial expansion reflects the brain's immediate adaptation to increased cognitive demands, such as managing lexical-semantic representations from multiple languages and engaging in language selection and inhibition. For instance, the left inferior parietal cortex often shows increased grey matter density that correlates with language proficiency and age of acquisition, indexing the brain's immediate plastic response to bilingual experience (Pliatsikas 460, 465).

Following this expansion phase, the DRM proposes a period of partial renormalisation or pruning, during which some previously expanded brain structures return toward baseline volumes as the bilingual individual develops greater efficiency in language processing and control. Rather than implying a loss of function, this renormalisation reflects an optimization and refinement of the underlying neural architecture. This dynamic pattern mirrors broader neuroplasticity principles observed in other complex skill-learning domains, where initial structural growth is followed by selective refinement for long-term efficiency (Pliatsikas 464).

Crucially, the DRM emphasises that these structural changes are not uniform across individuals; instead, they are modulated by the quantity and quality of bilingual language experiences. Factors such as the intensity of language use, switching frequency, proficiency, and age of acquisition all play a role. Variation along these experiential dimensions explains why some studies report volumetric increases while others find decreases or no change at all, as participants may sit at completely different points along the dynamic restructuring trajectory (Pliatsikas 465-466).

Subcortical structures implicated in language control, such as the basal ganglia and thalamus, similarly display the nonlinear volumetric adaptations predicted by the DRM. Volumes of the caudate nucleus and nucleus accumbens, for example, typically increase with early bilingual experience but can plateau or slightly decrease with very high levels of proficiency or extensive language-switching practice (Pliatsikas 462-465; Korenar et al. 1-3, 8-10). These adjustments in both grey and white matter highlight the continuous, bidirectional interaction

between structural neural changes and the functional demands of a bilingual lifestyle (Pliatsikas 466).

The non-linear trajectory proposed by this model introduces a theoretical paradox concerning structural expansion and renormalisation. Existing structural neuroimaging studies report diametrically opposed effect sizes for the exact same subcortical regions, displaying both volumetric expansion and volumetric reduction depending on the sample's profile. Without predefined, quantitative temporal benchmarks to distinguish the "expansion phase" from the "renormalisation phase," the framework risks becoming unfalsifiable, as any directional structural change—whether an increase, decrease, or null finding—can be interpreted as evidence of ongoing dynamic restructuring. This ambiguity complicates model testing. Clear timeline metrics remain urgently needed.

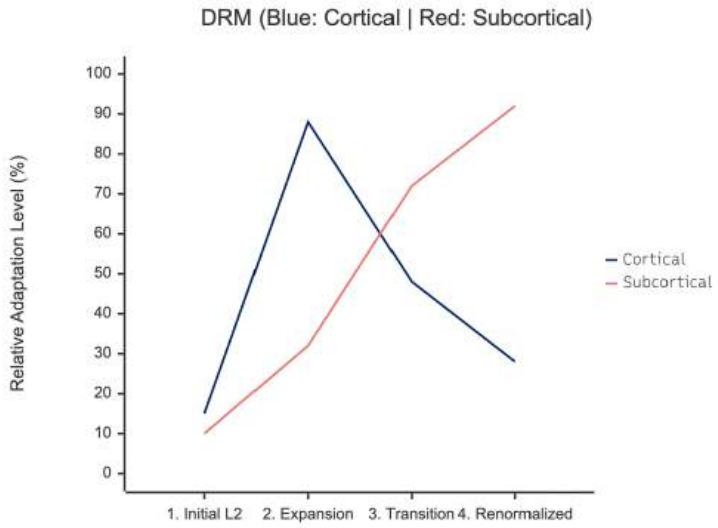


Figure 2: Dynamic Restructuring Trajectory. The dark blue line represents prefrontal cortical volume and energy-intensive top-down control engagement, illustrating rapid structural expansion during early L2 acquisition followed by partial renormalization (selective pruning) as control becomes more efficient. The dark red line tracks subcortical optimization (e.g., basal ganglia, striatum) and white-matter tract refinement, which steadily increases to support automated, low-cost language gating in balanced bilinguals. Developmental stages progress from (1) Initial L2 Exposure to (2) Early/Dominant Expansion Phase, (3) Intermediate Transition, and (4) Balanced/Renormalized Automated Processing.

Experimental Study

Aim

To investigate whether dominant and balanced bilinguals exhibit different multitasking performance.

Hypotheses

If continual reliance on executive control strengthens domain-general control systems, balanced bilinguals should outperform dominant bilinguals on multitasking tasks.

Alternatively, if balanced bilinguals have developed more efficient control networks through long-term adaptation and automatisation, they should perform equally well or better while expending fewer cognitive resources.

H1: Balanced bilinguals will outperform dominant bilinguals in multitasking.

H2: Dominant bilinguals will outperform balanced bilinguals in multitasking.

H0: Balanced bilinguals and dominant bilinguals will show negligible difference in multitasking performance.

Participants

Participants were recruited from a range of different settings:

- 32 participants from an international sixth form in the United Kingdom, where most students were bilingual.
- 40 participants from a multilingual city in Spain, where most of them speak Catalan (or other dialects) alongside Spanish.

The participants are separated into two non-manipulated groups: dominant bilinguals and balanced bilinguals. The sample comprised 36 dominant bilinguals, 21 males and 15 females, mean age of 19.3 (SD = 6.3). Participants were required to report at least three years of exposure to a second language. This criterion excluded very recent learners but did not by itself establish a stable language-dominance profile.

Where possible, balanced and dominant bilinguals were matched on age, sex, educational background, and language pair. For example, a dominant bilingual whose languages were English and Cantonese was matched with a balanced bilingual reporting the same language pair. The mean age is 20.2 (SD = 5.6), consisting of 21 males and 15 females.

Participants were recruited primarily through volunteer sampling, with some opportunity sampling. Some participants received financial compensation, whereas others participated without compensation. Participants below the legal age of consent provided assent and parental or guardian consent, whereas adult participants provided informed consent independently.

All participants were right-handed, had completed at least a secondary high school level of formal education, and were at least 16 years old. Participants reported no history of learning disabilities, social or communication impairments, uncorrected visual impairments, hearing impairments, or other conditions relevant to the study.

Materials

Participants completed an informed-consent form and a language-background questionnaire before beginning the task. Participants reported their age, educational background, handedness, and language history. Identifying information was stored separately from task data, and only information necessary for the study was collected. Later on, this data was deleted completely to keep anonymity of the participants.

An additional form for those under 18 or in high school (even if they are 18 at the time) were given to parents or guardians for formal consent. A reminder was given if the consent form is not filled within one week. If the consent form is not filled after two weeks or consent is denied, the participants would not take part in the study.

Participants completed a language-background questionnaire adapted from the Language Experience and Proficiency Questionnaire (LEAP-Q). They rated speaking, listening, reading, and writing proficiency in each language on a 1–6 scale, where 1 indicated very limited proficiency and 6 indicated very high proficiency. The questionnaire also recorded age of acquisition, current language use, years of formal instruction, and frequency of language switching.

Participants completed an Alternating Runs Task on the same laptop model in a standardised testing room. The task instructions, stimuli, response keys, timing, and trial sequence were identical for all participants. Link to the task used: Task switching

Procedures

Participants were informed that they could withdraw at any time without giving a reason and without penalty. All participation is fully voluntary and information regarding their name, age, gender, handedness, education and disability background has been deleted.

Participants were informed that the study examined language background and performance on a cognitive task. The specific prediction concerning language dominance and task switching was disclosed during debriefing.

Balanced Bilinguals were classified those whose two languages' proficiency are 4 or above; for example, someone with an L1 at 6 and an L2 at 4 were considered a balanced bilingual.

Dominant Bilinguals were classified those whose L1 proficiency is 4 or above and L2 is 3 or below; for example, someone with an L1 at 5 and an L2 at 2 were considered a dominant bilingual.

They then proceeded to do the Alternating Runs Task on our equipment in one of our standardised rooms (a classroom with white walls) and the researcher is expected to leave before the participant starts the task. After completion the data is gathered and every other irrelevant data was deleted. If the participants choose to withdraw from the study all data gathered from the participant was destroyed.

Data acquisition

The data acquisition started the 21st June 2026 and ended the 1st August 2026

The participants underwent an Alternating Runs Task in which their reaction seconds to the correct answers were recorded. Incorrect responses were excluded from reaction-time calculations but retained for the accuracy analysis. Trials continued until each participant completed at least 25 correct switch trials and 25 correct repeat trials.

$SC = \mu RT(switch) - \mu RT(repeat)$

Where μRT(switch) is the mean reaction time of correct trials where the task rule changed from the previous trial, and μRT(repeat) is the mean reaction time of correct trials where the task rule remained the same.

Accuracy was calculated as the percentage of correct responses separately for switch and repeat trials.

Before analysis, incorrect trials, anticipatory responses below 400 ms, and extremely slow responses above 1600 ms were excluded.

Findings

	Mean Repeat RT (ms)	Mean Switch RT (ms)	Mean Switch Cost RT (ms)
Balanced Bilinguals	739.4	877.3	137.9
Dominant Bilinguals	777.4	966.1	188.7

	SD Repeat RT (ms)	SD Switch RT (ms)	SD Switch Cost RT (ms)
Balanced Bilinguals	±155.1	±132.9	±45.7
Dominant Bilinguals	±215.8	±247.2	±67.5

	Accuracy Repeat	Accuracy Switch
Balanced Bilinguals	96.7%	96.2%
Dominant Bilinguals	96.3%	95.6%

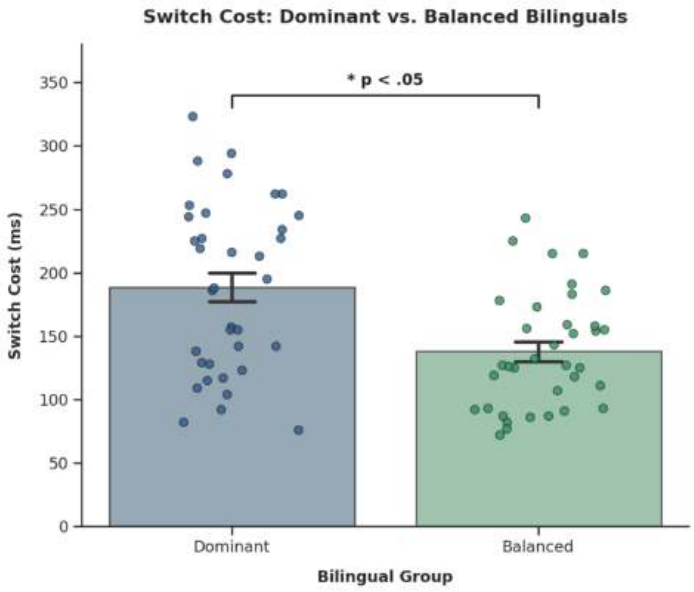


Figure 1. Comparison of Language Switch Cost (ms) Between Dominant and Balanced Bilinguals. Bar plots display mean switch costs with individual participant data points superimposed. Balanced bilinguals exhibited a significantly lower switch cost compared to dominant bilinguals.

A Mann-Whitney U test revealed a significant result, $U = 363.00$, $p < .05$. These results are based on the Switch Costs.

Balanced Bilinguals:

N = 36; Rank Sum = 1029.00

Dominant Bilinguals:

N = 36; Rank Sum = 1599.00

U-Test Results:

$U = 363.00$

$Z = -3.2102$

$p < 0.05$ (two-tailed)

$r \approx 0.38$

Discussion

The present findings align closely with current neurostructural framework models of bilingual language processing. Specifically, the observed response time advantages and reduced switch cost in balanced bilinguals is compatible with the Bilingual Anterior to Posterior and Subcortical Shift (BAPSS) model (but it does not directly test the neural mechanisms), which posits that increased proficiency and dual-language immersion shift the neural burden of executive control from energy-intensive prefrontal regions to more automated posterior and subcortical structures (such as the basal ganglia and anterior cingulate cortex). Because balanced bilinguals routinely manage two active language systems without a dominant preference, their cognitive control network becomes highly optimised, allowing for faster task-set reconfiguration during switch trials.

Furthermore, these results reflect the predictions of the Dynamic Restructuring Model (DRM), which conceptualises bilingualism as a continuous, adaptive process where executive control networks dynamically restructure based on the specific demands of language usage. The lower switch cost in balanced bilinguals suggests that sustained, high-level management of two languages drives neuroplastic adaptations that enhance domain-general cognitive flexibility. Critically, because accuracy remained uniformly high across both groups (~96%), these efficiency gains cannot be attributed to a speed-accuracy trade-off, but rather represent a genuine enhancement in processing speed and attentional control resulting from long-term, balanced language usage.

Limitations

Several methodological limitations should be considered when interpreting the findings of the study.

Group assignment relied on self-reported proficiency scores from a LEAP-Q adaptation rather than objective, direct language assessment measures. Furthermore, the use of volunteer sampling across two distinct geographic settings introduces potential cultural, educational and environmental variations. Moreover, the cross-sectional design prevents drawing causal inferences regarding whether balanced language usage directly drives cognitive flexibility over time or whether individuals with higher baseline executive control are more apt to attain a balanced profile.

Crucial background characteristics–including age of acquisition (AoA), daily language context (daily dual-language use), socioeconomic status (SES), and language exposure history–were not fully controlled or matched across groups, which could independently influence executive function and task-switching performance.

Findings also rely on a single experimental paradigm (Alternating Runs Task). Executive control comprises multiple distinct sub-components, such as working memory tasks and inhibitory control, which may not show the same pattern of results across task domains.

Conclusions

Balanced bilinguals may demonstrate superior cognitive control and task-switching efficiency compared to dominant bilinguals, responding significantly faster on both repeat trials (739.4 ms vs. 777.4 ms) and switch trials (877.3 ms vs. 966.1 ms) while incurring a smaller switch cost (137.9 ms vs. 188.7 ms). Because both groups maintain nearly identical, high accuracy levels across conditions (ranging from 95.6% to 96.7%), these findings indicate that balanced bilinguals possess enhanced cognitive flexibility in reconfiguring task sets without making a speed-accuracy trade-off.

Future directions

To build upon these findings, future research should integrate neuroimaging techniques such as fMRI or EEG to directly test whether the observed behavioural efficiency in balanced bilinguals corresponds to reduced prefrontal cortex activation and increased subcortical engagement, as predicted by the BAPSS model. Additionally, longitudinal studies are needed to observe how dynamic network restructuring unfolds over time as dominant bilinguals shift toward balanced proficiency. Scholars should also explore whether these task-switching advantages extend to other domains of executive function, such as working memory updating and inhibitory control, while controlling for potential confounding variables like age of acquisition, socio-economic status, and daily communicative context.

Works Cited

Anderson, John A. E., et al. "Effects of Bilingualism on White Matter Integrity in Older Adults." NeuroImage, vol. 167, 2017, pp. 143–150. https://doi.org/10.1016/j.neuroimage.2017.11.038. Accessed 4 Nov. 2025.

Bedore, Lisa M., et al. "The Measure Matters: Language Dominance Profiles across Measures in Spanish–English Bilingual Children." Bilingualism: Language and Cognition, vol. 15, no. 3, 2012, pp. 616–629. https://doi.org/10.1017/S1366728912000090. Accessed 23 Jan. 2026.

Calabria, Marco, et al. "Neural Basis of Bilingual Language Control." Annals of the New York Academy of Sciences, vol. 1426, no. 1, 2018, pp. 221–235. https://doi.org/10.1111/nyas.13879. Accessed 28 Apr. 2026.

DeLuca, Vincent, et al. "Understanding Bilingual Brain Function and Structure Changes? U Bet! A Unified Bilingual Experience Trajectory Model." Journal of Neurolinguistics, vol. 56, 2020, pp. 1–15. https://doi.org/10.1016/j.jneuroling.2020.100930. Accessed 29 Apr. 2026.

Genesee, Fred, et al. "Language Differentiation in Early Bilingual Development." Journal of Child Language, vol. 22, no. 3, 1995, pp. 611–631. https://doi.org/10.1017/S0305000900009971. Accessed 9 Dec. 2025.

Green, David, and Jubin Abutalebi. "Language Control in Bilinguals: The Adaptive Control Hypothesis." Journal of Cognitive Psychology, vol. 25, no. 5, 2013, pp. 515–530. https://doi.org/10.1080/20445911.2013.796377. Accessed 29 Mar. 2026.

Grundy, John G., et al. "Neural Correlates of Cognitive Processing in Monolinguals and Bilinguals." Annals of the New York Academy of Sciences, vol. 1396, no. 1, 2017, pp. 183-201. https://doi.org/10.1111/nyas.13333. Accessed 27 Apr. 2026.

Korenar, M., et al. "Bilingual Experiences Induce Dynamic Structural Changes to Basal Ganglia and the Thalamus." bioRxiv, 2021, pp. 1–10. https://doi.org/10.21203/rs.3.rs-1017465/v1. Accessed 30 June. 2026.

Lai, Gabrielle, and Beth O'Brien. "Examining Language Switching and Cognitive Control Through the Adaptive Control Hypothesis." Frontiers in Psychology, vol. 11, 2020, art. 1171. https://doi.org/10.3389/fpsyg.2020.01171. Accessed 13 Mar. 2026.

Lehtonen, Minna, et al. "Is Bilingualism Associated with Enhanced Executive Functioning in Adults? A Meta-Analytic Review." Psychological Bulletin, vol. 144, no. 4, 2018, pp. 394–425. https://doi.org/10.1037/bul0000142. Accessed 19 Oct. 2025.

Ng, Gee, and Hwajin Yang. "Code-Switching Patterns Differentially Shape Cognitive Control: Testing the Predictions of the Adaptive Control Hypothesis." Bilingualism: Language and Cognition, vol. 25, no. 3, 2021, pp. 521–535. https://doi.org/10.1017/s1366728921000754. Accessed 10 Apr. 2026.

Olsen, R., et al. "The Effect of Lifelong Bilingualism on Regional Grey and White Matter Volume." Brain Research, vol. 1612, 2015, pp. 128–139. https://doi.org/10.1016/j.brainres.2015.02.034. Accessed 9 Oct. 2025.

Oppenheim, Gary M., et al. "Longitudinal Evidence for Simultaneous Bilingual Language Development with Shifting Language Dominance, and How to Explain It." Language Learning, vol. 70, no. S2, 2020, pp. 20–44. https://doi.org/10.1111/lang.12398. Accessed 29 Jan. 2026.

Pliatsikas, Christos, et al. "The Effect of Bilingualism on Brain Development from Early Childhood to Young Adulthood." Brain Structure and Function, vol. 225, no. 7, 2020, pp. 2131–2152. https://doi.org/10.1007/s00429-020-02115-5. Accessed 18 Oct. 2025.

Pliatsikas, Christos. "Understanding Structural Plasticity in the Bilingual Brain: The Dynamic Restructuring Model." Bilingualism: Language and Cognition, vol. 23, no. 2, 2020, pp. 459-471. https://doi.org/10.1017/S1366728919000130. Accessed 03 Apr. 2026.

Stocco, Andrea, et al. "Conditional Routing of Information to the Cortex: A Model of the Basal Ganglia's Role in Cognitive Coordination." Psychological Review, vol. 117, no. 2, 2010, pp. 541–574. https://doi.org/10.1037/a0019077. Accessed 13 Apr. 2026.

Treffers-Daller, Jeanine. "What Defines Language Dominance in Bilinguals?" Annual Review of Linguistics, vol. 5, no. 1, 2019, pp. 375–393. https://doi.org/10.1146/annurev-linguistics-011817-045554. Accessed 19 Nov. 2025.

Unsworth, Sharon. "Amount of Exposure as a Proxy for Dominance in Bilingual Language Acquisition." Language Dominance in Bilinguals: Issues of Measurement and Operationalization, edited by Carmen Silva-Corvalán and Jeanine Treffers-Daller, Cambridge UP, 2015, pp. 156–173. https://doi.org/10.1017/CBO9781107375345.008. Accessed 24 Jan. 2026.

Wang Tao, et al. "An Activation Likelihood Estimation Meta-Analysis of How Language Balance Impacts the Neural Basis of Bilingual Language Control." Brain Sciences, vol. 15, no. 8, 2025, art. 803. https://doi.org/10.3390/brainsci15080803. Accessed 17 Mar. 2026.

Yamasaki, Brianna L., et al. "Effects of Bilingual Language Experience on Basal Ganglia Computations: A Dynamic Causal Modeling Test of the Conditional Routing Model." Brain and Language, vol. 197, 2019, art. 104665. https://doi.org/10.1016/j.bandl.2019.104665. Accessed 30 Apr. 2026.

Yee, Jon-Chao, et al. "Subcortical Restructuring as a Function of Multilingualism: Insights from Monolinguals, Bilinguals, Trilinguals and Quadrilinguals." Bilingualism: Language and Cognition, vol. 27, no. 1, 2024, pp. 715-728. https://doi.org/10.1017/S136672892300086X. Accessed 02 May. 2026.

Foreign Policy in Real Life:
How Teenagers Perceive the Relevance of Global Issues to Their Daily Lives

By Emma Orford

Abstract

Government officials rarely consider the impact on teenagers of foreign policy decisions. Nor do such officials typically try to persuade teenagers of the wisdom of their global policies. But while foreign policymakers might not focus on teenagers, their decisions often have a major impact on the teenage population. In this context, it is important to ask: 1) how teenagers perceive the relevance of foreign policy to their daily lives, 2) which foreign policy issues feel closest, and which feel most distant, and 3) whether teenagers believe they have political influence (even if they rarely do)? This study aimed to answer these three questions and to thereby add to our understanding of the potential mismatch between how policymakers view the relevance of the teenage population and how the teenage population views the importance of foreign policy decisions.

The study used a convergent mixed-methods approach, combining a Google Forms survey with interviews conducted by the researcher, both geared toward teenagers aged 13-19. Drawing on both methods, the study found that teenagers are most concerned with the impact of foreign policies on trade and prices, and are least concerned with issues of foreign aid and arms control. More generally, teenagers expressed strong interest in the topic of foreign policy, learned about such issues generally through social media, but had little expectation that their views had any impact. In addition, the interview component of the study produced six themes: foreign policy felt abstract until it affected the participant's family; economic effects were the most widely felt channel; proximity shaped salience; low efficacy coexisted with a sense of local agency; social media served as the primary window, with ambivalence; and witnessing issues firsthand drove engagement. These findings suggest that teenagers often feel unaffected by foreign policy until they directly experience it, and that they feel a low sense of power and efficacy regarding these issues. In sum, the study provides an exploratory baseline for future research into the intersection of teenagers' lives and foreign policy.

Keywords

adolescents, foreign policy, political socialization, political efficacy, youth civic engagement, mixed-methods research

Introduction

Every day, decisions made in rooms teenagers will never enter shape the texture of their lives. A tariff negotiated between governments changes the price of the clothes a 16-year-old can afford; a visa rule determines whether she sees her grandmother; a climate agreement (or its failure) reshapes the world she will inherit. Yet young people are rarely asked how foreign policy reaches them, and almost never included in making it. This disconnect - between those most

affected by policy and those consulted about it - raises a question that has received surprisingly little direct attention: how do teenagers themselves perceive the relevance of foreign policy to their own lives?

A substantial body of research helps frame this question, though none answers it directly. Three issues are especially relevant: how young people form political views, how they engage with civic and political life, and how their sense of political efficacy shapes that engagement.

First, research on political socialization establishes that political attitudes form early, well before voting age, and are shaped heavily by family, school, and the surrounding information environment (Jennings and Niemi; Sapiro). Adolescence is therefore not a political blank slate but a formative period in which views about the wider world are already taking shape (Neundorf and Smets; Flanagan). This matters for the present study because it suggests teenagers already hold perceptions of foreign affairs worth examining. They are not too young to have formed a relationship to these issues, even if that relationship is rarely studied directly.

Second, research on youth civic engagement shows that adolescents are more politically capable and engaged than they are given credit for and that their engagement increasingly runs through digital and social media rather than traditional civic institutions (Flanagan and Levine; Boulianne and Theocharis). This is directly relevant to how teenagers encounter foreign policy in the first place: if young people now learn about the world primarily through social platforms, then the channel through which foreign policy reaches them may differ sharply from the channels that shaped previous generations.

Third, research on political efficacy, a person's belief in their own capacity to understand and influence political processes, indicates that this sense of agency is a central predictor of whether young people engage with political issues at all (Niemi et al.). A teenager who believes she cannot affect a decision may disengage from it regardless of how much it affects her. Perceived relevance and perceived efficacy are thus intertwined: how close an issue feels and how much control one feels over it may together shape whether young people attend to it.

Finally, a small number of studies have examined how young people across countries relate to specific global issues, most notably large international surveys of youth attitudes toward climate change and its governance and studies focusing on how teenagers view conflict and peace resolution as well as their attitudes toward military service and wartime drafting (Hickman et al.; Covell et al.). Such work demonstrates both the feasibility and the value of asking teenagers directly about global issues at scale. The present study builds on that approach but broadens it: rather than a single issue, it examines how teenagers weigh the relevance of a range of foreign policy domains against one another.

Taken together, this literature establishes that teenagers form political views early, engage with the world increasingly through digital media, and act on issues in proportion to their sense of efficacy. What remains far less understood is how teenagers perceive the personal relevance of foreign policy, specifically, which international issues feel close to their daily lives and which feel distant, how those perceptions form, and whether they are consistent across national contexts. Existing research has largely examined domestic political attitudes, single global

issues, or civic engagement in general; the felt, everyday relevance of foreign policy to teenagers, in their own words and across countries, has received little direct attention.

The present study addresses this gap through a convergent mixed-methods design combining an international survey of teenagers (N = 75, across 21 countries) with a small set of in-depth interviews. It asks: **How do teenagers perceive the relevance of foreign policy to their daily lives, which issues feel closest and most distant, how do those perceptions form, and how much influence do they believe they have?** By pairing broad survey patterns with the detail of individual accounts, the study aims to describe not only which foreign policy issues teenagers feel affected by, but why.

This paper first describes the mixed-methods approach and instruments, then presents the survey and interview findings and their points of convergence, and finally discusses what these patterns suggest about how young people relate to the wider world and how that relationship might be strengthened.

Methods

Research design

This study used a mixed-methods design, in which both quantitative and qualitative data were collected and then compared for convergence (Creswell and Plano Clark). An online survey measured how strongly teenagers perceived a range of foreign policy issues to affect their daily lives, while a small set of semi-structured interviews explored, in participants' own words, how and why those issues became personally relevant. The survey was designed to capture the broad pattern across many respondents; the interviews were designed to add explanatory depth. The two strands were analyzed separately and then integrated to assess whether they pointed to the same conclusions.

Participants and recruitment

Participants were teenagers aged 13-19. They were recruited through the researcher's personal and extended networks and through snowball sampling, in which initial participants were asked to share the survey with peers, including peers in other countries. Recruitment also occurred through the social media presence of a companion podcast produced by the researcher on the same topic. This was a convenience sample and is not representative of teenagers generally; its purpose was exploratory.

The survey received 75 valid responses from teenagers in 21 countries across six world regions: North America (n = 40), Europe (n = 20), Africa (n = 7), Oceania (n = 4), Asia (n = 3), and the Middle East (n = 1). Respondents' ages ranged across 13-19, with 17 being the most common. Four teenagers took part in interviews; all were recruited through the researcher's local and personal network and were based primarily in the United States, though several had international family backgrounds or prior experience living abroad. Because the interview

sample was small and geographically concentrated, interview findings are treated as illustrative rather than representative.

Instruments

The survey contained eight items and took approximately three minutes to complete. After answering two demographic items (age and country), respondents were asked to rate, on a five-point scale from 1 ("Not at all") to 5 ("A lot"), how much each of 10 foreign policy areas had directly affected them or their family: visa and travel rules; trade and prices; climate and environmental agreements; immigration and refugee policies; conflict and safety; internet and technology rules/cybersecurity; health care policy/public health; human rights advocacy; foreign aid and development; and arms control. Respondents further answered three questions made up of a five-point agreement scale (perceived ability of young people to influence foreign policy decisions; perceived understanding of how foreign policy connects to daily life; and desire to learn more). Additional questions captured where respondents typically hear about international issues (multiple-select) and which issues they were most interested in learning about. An optional open response item invited a short anonymous account of any personal impact. The full instrument appears in Appendix A.

For the interviewees, an interview guide was used. Each interview followed a standardized core block of open-ended questions asked in the same order of every participant. These questions covered a brief warm up question, baseline perceptions of foreign policy, perceived relevance to daily life, any specific experience of a government or international decision affecting the participant, a structured ranking exercise, sources of information, if the interviewee thoughts they could affect foreign policy decisions, engagement drivers, and a closing question. In the ranking exercise, the interviewer read the same list of ten foreign policy areas, in the same order, to every participant, and recorded which single area each participant identified as closest to their daily life and which as most distant, along with their reasoning. The full interview guide appears in Appendix B.

Data collection

The survey was administered online via Google Forms and remained open from July 14, 2026 to August 6, 2026. It was anonymous: no names, emails, or identifying information were collected, apart from an optional, separate invitation for respondents interested in appearing on the companion podcast, which was not linked to survey responses used in analysis. Interviews were conducted individually over video calls on Riverside and took approximately 10 minutes each. They were audio recorded with consent and transcribed. Interview recordings used for research were the full, unedited sessions. Any later edited audio produced for the podcast was a separate product and was not used as research data.

Analysis

Survey responses were analyzed with descriptive statistics. For each of the 10 policy

areas, an average impact rating was calculated across all respondents, and areas were ranked to identify which felt most and least proximate to teenagers' lives. The percent of respondents reporting meaningful impact (a rating of 3 or higher) and no impact (a rating of 1) was also calculated for each area, along with the average agreement for each attitude item and the distribution of information sources.

Interview transcripts were analyzed using thematic analysis (Braun and Clarke). The researcher first read all transcripts without coding to gain familiarity, then using AI support developed a codebook of recurring themes, then coded each transcript against that codebook in a shared matrix, recording supporting excerpts. Because a single researcher conducted the coding, the transcripts were re-read after an interval to check for consistency of coding. The two data strands were then compared to identify points of convergence and divergence. The final codebook appears in Appendix C.

Ethics and disclosures

The study was conducted as independent student research under the informal guidance of a faculty advisor in health policy at Columbia University. As independent student research, it was not conducted under, and was not eligible for, formal institutional review board oversight. In place of such review, the interview guide, survey, and consent form (Appendix D) were reviewed by the faculty advisor before data collection. Because interview participants were minors, written consent was obtained from each participant and written permission from a parent or guardian before any interview. Interview participants are identified in this paper only by age and country (for example, “a 17-year-old in the United States”). A private key linking pseudonyms to identities was retained separately by the researcher and is not published. Survey participation was anonymous, with an opening consent statement and the survey was open to teenagers aged 13-19.

Use of AI tools and other software

The researcher used AI-based tools to support several stages of this project: transcribing interview audio, performing an initial pass of survey computation, generating a first pass of thematic coding, and assisting with structuring, drafting, and editing the manuscript. The researcher reviewed, verified, corrected, and finalized all of this work, and all analytical decisions, interpretations, and conclusions are the researcher's own. The researcher also used Zotero to help collect, manage, organize, and cite the research sources.

Results

Which issues feel closest (survey)

Across the 75 respondents, the foreign policy areas rated as having the greatest direct impact on daily life were the everyday and economic ones, while more abstract geopolitical categories were rated lowest. Table 1 presents the average impact rating for each area. Trade and

prices received the highest average rating (2.53), followed by health care / public health (2.39) and visa and travel rules (2.23). At the other end, foreign aid and development (1.71), arms control (1.72), immigration and refugee policies (1.77), and climate and environmental agreements (1.77) were rated as least impactful. No area exceeded a mean of 2.53 on the five-point scale, indicating that most teenagers perceived the direct impact of foreign policy on their lives as modest overall, even for the highest-rated issues.

Rank	Policy Area	Average (1-5)	% Rating ≥3	% Rating 1
1	Trade and prices	2.53	47	24
2	Health care / public health	2.39	43	32
3	Visa and travel rules	2.23	39	49
4	Conflict and safety	2.15	35	47
5	Internet and tech / cybersecurity	2.01	32	48
6	Human rights advocacy	1.92	25	53
7	Climate and environmental agreements	1.77	21	48
7	Immigration and refugee policies	1.77	19	56
9	Arms control	1.72	17	59
10	Foreign aid and development	1.71	20	64

Table 1: Average perceived impact of each foreign policy area on daily life (1 = not at all, 5 = a lot; N = 75), ranked.

Attitudes: high interest, low efficacy (survey)

Respondents expressed strong interest in the topic but low confidence in their ability to shape it. The average agreement was highest for “I want to learn more about these issues” (M = 3.75) and moderate for “I understand how foreign policy connects to my life” (M = 3.27), but lowest for “People my age can influence foreign policy decisions” (M = 2.84) - the only attitude item to fall below the scale midpoint. This gap between interest and perceived influence was one of the clearest patterns in the survey. However, the average was still relatively low, showcasing the disconnect many teenagers feel regarding foreign policy.

Information sources (survey)

Social media was by far the most common way respondents encountered international issues, reported by 84% of the sample, ahead of news media (68%), family (53%), school (45%), and friends (39%). Only 8% reported that they rarely hear about international issues. When asked which issues they most wanted to learn about, respondents most often selected conflict and safety, human rights advocacy, and trade and prices. Notably, the issues of greatest interest were not the same as those of greatest perceived impact.

Interview themes (qualitative)

Thematic analysis of the four interviews produced six themes, summarized here with illustrative excerpts.

Theme 1: Foreign policy felt abstract until it affected the participant's family. All four participants initially framed foreign policy as remote before describing a concrete personal or family impact. One participant, whose parents worked in international development, described foreign policy as "stuff overseas… that doesn't really affect us," and then, in the same interview, explained that the closure of a foreign aid agency had eliminated both parents' jobs. Another, of Ukrainian heritage, described losing family members following the invasion of Ukraine.

Theme 2: Economic effects were the most widely felt channel. Even participants without a dramatic personal story pointed to prices and tariffs as the point where foreign policy touched their lives, citing higher costs for imported goods, aligning with trade and prices ranking highest in the survey.

Theme 3: Proximity shaped salience. The issue each participant identified as "closest" mapped directly onto their own circumstances: foreign aid for the child of aid workers, conflict and immigration for the participant with family abroad, arms control for a participant who described fearing for safety at school, and health care and technology for a participant focused on community needs.

Theme 4: Low efficacy alongside a sense of local agency. Participants doubted they could influence decisions directly but several distinguished between influencing decisions and influencing local impact. One observed that teenagers cannot do much "to influence these decisions, but… they can influence the impact it has… help their community", echoing the survey's low efficacy score.

Theme 5: Social media as the primary window, with ambivalence. Consistent with the survey, participants named social media as a main source of information while noting its limits; one described it as offering "a headline you can understand" even if it "doesn't tell you the truth."

Theme 6: Witnessing issues firsthand drove engagement. Direct observation, rather than news coverage, motivated concern. One participant described seeing the loss of glacier ice while traveling and contrasting it with climate agreements not being met.

Convergence of survey and interviews

The two methods and results converged on the study's central pattern: teenagers perceived foreign policy as relevant in proportion to their direct, personal contact with it. The survey showed that the most everyday, widely encountered issues (trade and prices, health care) were rated most impactful, while abstract categories (foreign aid, arms control) ranked lowest; the interviews explained why, with each participant's sense of relevance rooted in a specific lived experience. The strands also converged on two secondary findings: the dominance of social media as an information source (84% of survey respondents; three of four interviewees), and the gap between interest and efficacy (survey means of 3.75 versus 2.84; articulated by interviewees who wanted to be involved but doubted their influence). One divergence emerged: although trade

and prices was rated most impactful in the survey, it was not among the issues teenagers were most interested in. One interviewee ranked it the most distant from his life despite acknowledging its effects, suggesting that felt impact and active interest are distinct issues.

Discussion

This study found that proximity shapes salience: teenagers perceive foreign policy as most relevant when they have direct community or familial ties to an issue. Direct witness and experience drive the felt impact and perceived relevance of foreign policy decisions. The study also found that the everyday beats the geopolitical: the issues that feel closest and most personal are often the opposite of what dominates foreign policy headlines. In addition, teenagers expressed a strong sense of local agency despite low political efficacy. As seen in the interviews, many teenagers felt they could not affect decision-making at the national or international level, but believed they could still make a local impact, such as a tangible change in their immediate community. Finally, the study found that social media is the dominant source of information on these issues, and that interest is not synonymous with impact when it comes to foreign policy.

The researcher believes that trade and prices and health care felt closest, while arms control and foreign aid felt distant, simply because of what is most present in teenagers' lives: an issue feels close when someone has direct experience with it. Because a large share of respondents and interviewees were based in North America, and more specifically the United States, it is logical that trade and health felt close. The United States health care system has its pros and cons, but everyone interacts with it in some way, whether that means finding the money to pay a hospital bill for a relative fighting cancer or appreciating easy access to an x-ray machine after breaking an arm. Foreign aid, on the other hand, likely felt distant because the United States is an economically developed country that does not only rely on foreign aid to provide for its citizens. Notably, the survey results showed that the respondents who did feel impacted by foreign aid tended to be based outside of the United States, or to have family or friends who had worked globally and been directly affected by it.

In the interviews, what most often changed an issue from abstract to real was, again, direct personal experience: a parent losing a job or a fear of school shootings. This offers valuable insight into how teenagers come to care about global issues. It is interesting that even when a teenager was not especially aware of events on a global scale, they were still aware of how specific policies could reach them. For example, the interviewee who described her fear of school shootings also shared that she had never traveled abroad and did not feel very affected by foreign policy, yet she recognized that her daily life was shaped by arms control.

Teenagers also shared, through the survey, that they are interested in learning more about foreign policy but do not feel they can influence it. The researcher believes this reflects realism rather than apathy. At the national or international level, it is difficult for a single teenager to make a large impact on policy decisions, yet locally the chances for creating change are much higher. This sentiment appeared in both the survey and interview responses. Because teenagers

seemed to redirect their agency to the local level, it appears that this is not teenager apathy, but it is teenagers reallocating effort toward where impact feels realistic.

Although this research did not set out to examine it directly, the results were consistent with previous work on political socialization (Jennings and Niemi; Sapiro). The idea that teens' political beliefs are shaped early by their communities and environments appeared in the survey's short-answer question, which asked respondents to elaborate on how foreign policy had affected them. One respondent wrote, “My family is very open on the fact that they love guns and do not want them restricted.” This teenager named a familial belief and connected it to their own views on foreign policy and its perceived impact, aligning with prior claims that the political beliefs a teenager grows up around shape their own.

This study also extends prior work on youth civic engagement and youth digital political activity by showing that the number one resource for learning about international issues among teenagers is social media. The survey provides concrete evidence: 84% of teenagers gathered their information on foreign policy and international issues from social media. Given that they gather most of their information there, it is logical that they would also be more politically active on similar digital platforms than through physical or traditional media. One respondent wrote, “Posting about human rights on the Internet has gotten my account shut down,” corroborating prior work suggesting that teenagers are politically active in digital forms and are, to an extent, aware of current affairs (Boulianne and Theocharis).

The participants in this study also showed ideas that slightly differ from previous work on political efficacy. Prior research suggests that how close an issue feels and how much control someone believes they have over it will shape how they engage with it (especially for youth) and that feeling less control leads to less engagement (Niemi et al.). The data from this study suggest that teenagers doubt their ability to influence large international decisions but carry a strong belief in local impact. This both reinforces and refines that idea: these teenagers feel they have little impact on international decisions and therefore rarely engage with them directly, yet in nearly all communities teenagers are striving to make local impacts, whether by leading fundraisers, creating service projects, or simply reaching out to elected officials. One interviewee said, “There's really not a lot that I think teenagers my age can do to influence these decisions… but what I think they can do is influence the impact it has… help their community,” which aligns well with the finding that teenagers feel they can impact local issues.

The mixed-methods approach added a convergence point worth noting. The survey showed the broad numbers and patterns; the interviews gave context and a personal dimension, and the two aligned relatively well. For example, both represent a pattern of high interest and low sense of power: teenagers want to know more, but a clear efficacy gap remains, as they feel they cannot influence policy decisions themselves. The methods also converged on the finding that impact is not synonymous with interest, though they diverged on trade and prices. In the survey, trade and prices consistently ranked among the most impactful, yet was not near the top of what respondents wanted to learn more about (conflict and safety was the most frequently chosen topic there), and in the interviews, one participant ranked trade and prices as the most

distant issue from his life. This showcases the value of using two methods, as well as how the perceived impact of foreign policy differs based on individual lived experience. Even so, the survey and interviews produced similar results overall: perceived impact was closely tied to personal experience, social media was the dominant way teenagers gathered information, and there was a sizable perceived efficacy gap regarding international issues.

Using both methods allowed the researcher to collect data in different settings. This was important because the way someone answers an anonymous survey may differ from how they respond in an interview. The mixed-method approach also added depth, with the ability to see both the broader pattern and the individual experience behind it. The survey devoted a single question to briefly explaining how a foreign policy issue had affected the respondent, whereas the interviews were structured so that every question led toward individual experience and remained open-ended to allow a fuller conversation about how different issues and policies affected the individual. Using both methods also allowed each to validate the other and enabled cross-checking to test accuracy and better determine convergence and divergence. Had the interview stories contradicted the survey statistics, the researcher could have highlighted a nuanced gap that would not be possible to identify using only one method.

These findings can offer valuable insight into how global issues are taught and discussed with teens. Relevance appears to stem from lived experience and these findings suggest it is important that adults and teachers highlight not only why these issues matter, but also the direct and indirect impacts they have on a teen's life. Doing so could raise teenager awareness of global events and increase political involvement and civic engagement, ultimately, shifting perceived efficacy around both local and international issues. One possible implication is that the efficacy gap calls for a different approach to these topics: informing teenagers about the issues while also emphasizing the impact even a single person can have on foreign policy.

If this study were replicated with more resources and reach, future research could be strengthened in several ways, for example, by creating multilingual surveys to reach more teenagers around the world, and by adding interviews with teenagers from different countries, cultures, and socio-economic backgrounds to reduce the inherent bias of this exploratory research. The researcher also believes that adding survey questions about a respondent's socio-economic background and gender would add valuable information and weave in another dimension when analyzing perceived impact and efficacy across individual circumstances. This study also raised further questions worth exploring, such as whether the efficacy gap looks different in different countries. Overall, these findings set an exploratory baseline for future research and should be built on with a wider range of participants and a greater representation of teenagers around the world.

Limitations

Several limitations should be considered when interpreting these findings. Most stem from the exploratory, small-scale nature of the study and suggest directions for future research rather than undermining its central patterns.

Participants were recruited through the researcher's personal and extended networks and through snowball sampling, producing a convenience sample rather than a representative one. The survey sample, while spanning 21 countries, was concentrated in North America and Europe, and respondents were reached largely through the researcher's own social and online networks. Teenagers without access to those networks or without internet access could not participate. Findings describe patterns within this particular sample and should not be generalized to teenagers worldwide.

The qualitative strand consisted of four in-depth interviews and all interviewees were based in the United States, though several had international backgrounds. Four interviews cannot capture the full range of teenagers' experiences and the themes identified are illustrative rather than representative. Their purpose was to add depth and possible explanations to the survey patterns, not to establish independent generalizable claims, and they are best read as hypotheses for future, larger qualitative work.

The thematic analysis of the interviews was conducted by a single researcher, which introduces the possibility of individual bias in how themes were identified and applied. To partially address this, transcripts were re-read after an interval to check that coding was applied consistently, but the analysis did not include a second independent coder, which would strengthen future work.

The study measured teenagers' own perceptions (how much they felt each issue affected them, and how much influence they believed they had) rather than any objective measure of impact or engagement. Perceived relevance is the intended focus of the study, but self-reported perceptions can be shaped by awareness, recent events, or social desirability, and may not correspond to the actual degree to which a given policy affects a respondent's life.

The central survey item measured perceived impact on a five-point scale rather than separately distinguishing whether a respondent had experienced an issue from how strongly they rated its relevance; as a result, experience and perceived impact could not be analyzed as fully independent variables. In addition, the survey was administered only in English, which restricted participation to teenagers with English proficiency and likely contributed to the sample's regional concentration; translated instruments would broaden reach in future studies.

Finally, as an independent high school research project conducted over a single summer, the study was limited in scale, duration, and resources. It was not subject to formal institutional ethical review, though its instruments and consent procedures were reviewed by the researcher’s informal faculty advisor at Columbia University. A larger, multilingual survey, a broader and more geographically diverse interview sample, and multiple coders would each strengthen the reliability and reach of the conclusions.

Conclusion

In conclusion, this study aimed to provide exploratory findings regarding how teenagers perceive the relevance of foreign policy to their daily lives. Specifically, the study strove to find insight into which issues feel closest and which feel most distant, how the perceptions form, and

how much influence teenagers believe they have over these issues. The critical results of the study highlight the foreign policy perceptions of teenagers around the world: foreign policy felt abstract until it affected the participant directly and witnessing issues firsthand drove engagement drastically; proximity shaped salience; low efficacy coexisted with a sense of local agency; and social media served as teens' primary window into international issues. These findings imply that teenagers broadly feel unaffected by foreign policy until they are directly impacted and have first-hand experience, and that they are relatively politically aware. This study is important because it creates an exploratory baseline for future research to build upon, and it provides unique perspectives, given that the interviews were conducted by a teenager, creating more authentic and conversational feelings, yielding more organic and example-based results. However, it is important to note that this study had significant limitations, including a limited number of respondents, the majority of whom were based in the United States, and a necessary proficiency in English required to participate. In the future, this study could be enhanced by reaching more teenagers in countries from around the world, translating the survey into other languages, and creating more detailed survey questions to properly analyze individual circumstances and how that impacts the perceived relevance of foreign policy. To dive deeper into how teenagers perceive foreign policy in comparison to other age groups, a future study could also ask a similar set of questions to working age adults and seniors, and analyze the extent to which sentiments and perceived relevance of foreign policy differ by age group. By further exploring this area, researchers can better understand the impact foreign policy has on teenagers globally and ideate new ways to involve teenagers in the global doings of the world.

Acknowledgments

The researcher would like to thank Dr. Michael Sparer, William Henry Welch Professor of Public Health Systems and Director at the Center for Public Health Systems at Columbia University Mailman School of Public Health, for serving as an informal faculty advisor to this project and reviewing the survey instrument and providing feedback and guidance on the research. The analyses, interpretations, and any errors are the author's own.

Works Cited

Boulianne, Shelley, and Yannis Theocharis. “Young People, Digital Media, and Engagement: A Meta-Analysis of Research.” Social Science Computer Review, vol. 38, no. 2, Apr. 2020, pp. 111–27. SAGE Journals, https://doi.org/10.1177/0894439318814190.

Braun, Virginia, and Victoria Clarke. “Using Thematic Analysis in Psychology.” Qualitative Research in Psychology, vol. 3, no. 2, Jan. 2006, pp. 77–101. Taylor and Francis+NEJM, https://doi.org/10.1191/1478088706qp063oa.

Covell, Katherine, et al. “Age Differences in Understanding Peace, War, and Conflict Resolution.” International Journal of Behavioral Development, vol. 17, no. 4, Dec. 1994, pp. 717–37. SAGE Journals, https://doi.org/10.1177/016502549401700409.

Creswell, John, and Vicki Plano Clark. Designing and Conducting Mixed Methods Research. Sage Publications, 2006.

Flanagan, Constance. “Teenage Citizens.” Harvard University Press, https://www.hup.harvard.edu/books/9780674048621. Accessed 15 Aug. 2026.

Flanagan, Constance, and Peter Levine. “Civic Engagement and the Transition to Adulthood.” The Future of Children, Jan. 2010. www.academia.edu, https://doi.org/10.1353/FOC.0.0043.

Hickman, Caroline, et al. “Climate Anxiety in Children and Young People and Their Beliefs about Government Responses to Climate Change: A Global Survey.” The Lancet Planetary Health, vol. 5, no. 12, Dec. 2021, pp. e863–73. www.thelancet.com, https://doi.org/10.1016/S2542-5196(21)00278-3.

Jennings, M. Kent, and Richard G. Niemi. “The Transmission of Political Values from Parent to Child.” American Political Science Review, vol. 62, no. 1, Mar. 1968, pp. 169–84. Cambridge University Press, https://doi.org/10.2307/1953332.

Neundorf, Anja, and Kaat Smets. “Political Socialization and the Making of Citizens.” Oxford Handbook Topics in Politics, edited by Oxford Handbooks Editorial Board, Oxford University Press, p. 0. Silverchair, https://doi.org/10.1093/oxfordhb/9780199935307.013.98. Accessed 15 Aug. 2026.

Niemi, Richard G., et al. “Measuring Internal Political Efficacy in the 1988 National Election Study.” The American Political Science Review, vol. 85, no. 4, 1991, pp. 1407–13. JSTOR, https://doi.org/10.2307/1963953.

Sapiro, Virginia. “NOT YOUR PARENTS’ POLITICAL SOCIALIZATION: Introduction for a New Generation.” Annual Review of Political Science, vol. 7, 2004, pp. 1–23. Annual Reviews, https://doi.org/10.1146/annurev.polisci.7.012003.104840.

Virtue or Verdict: A comparative study of Sulh-i-Kul and Bhinneka Tunggal Ika on Legal Pluralism, Sovereignty and Historical Legacy By Prathama Mandal

A. Introduction

In 1580, inside the dark sandstone halls of Fatehpur Sikri, the Mughal Emperor Akbar sat listening to Hindu pandits, Jesuit priests and Muslim clerics debating on theology (Mallampalli). The debate was not meant to christen a single faith that the empire should follow, rather it was meant to prove that the mighty Mughal throne requires none. Similarly, nearly four thousand miles east and two centuries earlier, a Buddhist poet on the island of
Java inside the court of Majapahit named Mpu Tantular inscribed a single line based on a remarkably similar philosophy: "Bhinneka Tunggal Ika Tan Hana Dharma Mangrwa" – literally means, "Although it is divided, it is yet one. The truth is clear and unmistakable" (Museum Nasional Indonesia).

Intriguingly, both men were not drafting codified legislation. Akbar was trying to unify the frontiers of an empire fractured by religious sentiments, whereas, Tantular was busy reuniting Shiva-worship and Buddhist sects, both important factions of the royal court. In spite of that, both of them reached the same conclusion – pluralism as a form of governance instead of tolerance rooted in pity. It would be quite appealing to group them with a homogenous label – progressive individuals; visionaries who were ahead of their times. But this paper attempts to resist this temptation because Sulh-i-Kul and Bhinneka Tunggal Ika were not ideas that can be associated with a single dynastic ruling. They were proto-ideas, mostly imperfect measures devised to solve a problem that is still relevant in modern times – the idea of bringing people of differing sects and faiths under the single, unified umbrella of statehood.

B. Sulh-i-Kul: Harbour of diversity or Monarchical Ambition

Akbar's idea of Sulh-i-Kul was not just a fantasy, but a futuristic and liberal solution towards a problem that plagued the empire. The idea was actually moulded into lived realities of common people, rather than just being an abstract, ambitious concept of the emperor. For example, Ibadat Khana became a place of discussion and discourse on matters of theology, transforming plausible areas of conflicts into intellectually stimulating conversations.

Alongside, in 1564 Akbar abolished the jizya tax, a revenue system based on religious identity (Richards). Additionally, he also ensured impartial appointments of Qazis (judges) instead of favouring a particular religion (Munir). His ban on cow slaughtering was a deliberate incorporation of Hindu beliefs (Zehra and Singh). These showed that Sulh-i-Kul ensured pluralism was adopted as a form of governance and did not remain a mere self-righteous proclamation. The policy effectively united people of varying faiths, thereby making the empire a tolerant space, committed to inclusivity. And yet even this most comprehensive expression of Sulh-i-Kul is open to analysis. In 1582, Akbar systematized Din-i-Ilahi, a heterogeneous movement combining elements of Islam, Hinduism, Zoroastrianism and Jainism, with the king being the spiritual pivot (Richards).

However, historians remain divided on the significance of this act – some see it as a rational extension of the pluralistic vision of Sulh-i-Kul and a genuine effort to establish common ground between diverse faiths, whereas others believe the exact opposite – a monarch integrating the religious power with himself, rather than distributing it equally among his religiously diverse subjects. If the latter interpretation holds more weight, then Sulh-i-Kul was never about tolerance and was just an inseparable entity of Akbar's own model of sacred monarchy.

Besides, the very grandeur of the policy exposes its instability. This was because the formation and implementation of the radical reforms depended solely on the goodwill of one single person – the emperor. The idea became further strengthened with the introduction of Mazhar of 1579 – a decree that gave Akbar the power to overrule conflicting interpretations of Islamic law (Sharia), making him the sole custodian of pluralism. Needless to say, a policy that relies solely on the temperament of its creator, no matter how magnanimous, cannot, in its truest sense, be considered a law.

The fragility was no abstract concept. Aurangzeb's reign is the best example of this grim reality in stark detail: In 1679, a century after the dissolution of jizya tax policy, he reinstated the practice, consequently reversing Akbar's single act of fiscal tolerance (Richards). Temple policy also changed with some being destroyed or facing restriction during his reign (Eaton). There was no alternative structure that mimicked the openness of Ibadat Khana in court. All these were not consequences of new changes in the law, but all it required was a change in the disposition of the ruler; which is the exact vulnerability emphasized in this paper.

C. Bhinneka Tunggal Ika: Didactic Poem or Political Convenience

In contrast to Sulh-i-Kul which was unsuccessful due to its excessive dependence on a single authoritative source, i.e. the state, the Majapahit doctrine was unfortunately too detached from the notion of statehood from the very beginning. The verses of Mpu Tantular formed in the Kakawin Sutasoma were written initially for a different purpose. It strived for a specific goal – the reconciliation of Buddhist and Shaivite sects and ensuring common understanding that these are simply different forms of one single universal truth (Aghababian).

Indeed, while Sulh-i-Kul did not have such a fortune to be incorporated into modern legislation, Bhinneka Tunggal Ika had, impressively enough, gained official status as the national emblem of Indonesia under Article 36A of the 1945 constitution (Constitution of the Republic of Indonesia). This separation from statehood is further underscored by the working of the internal politics of Majapahit itself. The Mughal empire consists of a bureaucratic state that ruled a relatively united population (Richards). On the contrary, Majapahit was a mandala state: a central kingdom in East Java, encircled by tributary kingdoms whose autonomy increased with increasing distance from the centre (Wolters). There were no empire-wide legal mechanisms through which Bhinneka Tunggal could be validated as a statute. The fact that it was adopted as a legal document is unusually remarkable, since it managed to acquire the structure that Majapahit's governing structure failed to provide.

Even though this symbolises a strategic move, the transformation of a poetic notion to legislation is rather controversial. The initial conception of Bhinneka Tunggal Ika was

connected with the Buddhist and Hindu doctrine (Aghababian), however, with the spread of Islam in the archipelago, the idea was improvised to include a faith that was not accounted for in the original version (Ricklefs). This makes the idea an entity of state affairs instead of being a true harbinger of diversity.

Moreover, at times the Indonesian authorities have appealed to the heritage of Majapahit to justify claims of territorial and maritime ownership (Bloembergen and Eickhoff). This resulted in a genuine philosophical idea being transformed into a selfish measure of political gain. The fact that this measure can be used to legitimize both regional diversity and territorial annexations, leads to uncomfortable questions regarding its reliability and endurance. History has testified that the concept has been used by different individuals and communities in different ages, thus implying that it is a malleable idea that is open to diverse interpretations well suited for politically motivated gains.

The fact that it can be used to justify both pluralism and nationalism signifies that the philosophy lacks coherence and consistency. Though Bhinneka Tunggal Ika had the potential to be a powerful symbol of unity and diversity, its continued saga of interpretations and reinterpretations have weakened its stance.

D. The Problem that Outstays its Solutions

It is tempting to conclude that pluralism as a form of governance is impossible and is destined to become either an arbitrary rule of a ruler or convenient policies of the nation-state. However, this paper aims to avoid such a conclusion, instead it will argue that identified flaws indicate a single thesis: Pluralism cannot exist as individual virtue and must be an apt instrument of the state.

The asymmetry can be charted on two axes namely: form, that is whether the ideology achieves the status as a legally binding structure independent of its originator — and content — whether the essence of that ideology survives even after being codified. Sulh-i-Kul and Bhinneka Tunggal Ika fall at almost the exact opposite ends of the continuum.

Bhinneka Tunggal Ika in such a case can be a corroboration. A profound poetic philosophy composed by Mpu Tantular with the sole intention of reconciling two varying sects (Aghababian) – without any ambition of legislation – ironically became an indisputable part of Indonesian legislation and government regulations six centuries later. All this change did not occur in a single momentous act. The phrase ‘Bhinneka Tunggal Ika Tan Hana Dharma Mangrwa’ was inscribed into the national emblem – Garuda Pancasila – in 1950 (Government Regulation No. 66 of 1951), and Sukarno's independence movement adopted Bhinneka Tunggal Ika as its political ideology in 1945 (Binsasi et al.). The process of becoming law took decades, and only in 2011 did the phrase gain the status as a legal principle instead of being a mere ideology. The journey from 1945 to 2011 shows the gradual, conditional path of Bhinneka Tunggal Ika, a stark reminder that the line between law and belief is permeable and can be crossed by purposeful state actions.

Paradoxically, it also represents an absurd inversion of expectations stemming from histories of past institutions: Sulh-i-Kul was a policy rooted in pillars of governance – through

courts, judges, tax regulations (Richards) – whereas, Bhinneka Tunggal Ika was a doctrine that was severed from statehood entirely. Yet, it was the poem, not the policy, that outlived its era and got immortalised as codified legislation. The dogma with administrative framework died with its administrator; whilst a simple didactic poem armed with teeth exceeded all expectations of its creator. Sulh-i-Kul's tragedy was its association with Akbar's legacy, while Bhinneka Tunggal Ika's fortune lay in its disconnection from a single ruler, allowing it to be freely claimed by figures of the future.

Even though Bhinneka Tunggal Ika's afterlife is quite literal, Sulh-i-Kul's afterlife is far more consequential. Sulh-i-Kul did not have any defining moment of codification that could be identified as a manifestation of Akbar's policy in law. It has instead reappeared as a continuing legal issue that arises whenever a pluralistic society needs to determine the legal neutrality of the state concerning different religions. Articles 25-28 of the Indian Constitution, granting freedom of religion and ensuring non-interference of the state (Constitution of India), are not direct heirs of Sulh-i-Kul – otherwise it would be the very sort of historical analogy that this paper sought to avoid. But both attempt to solve the same problem faced by Akbar in Fatehpur Sikri – the process of establishing a neutral secular state in the absence of a sovereign capable of transcending all religious boundaries. Akbar depended on his own personhood; in contrast constitutional democracies achieved structural stability by integrating courts and fundamental rights that Sulh-i-Kul greatly lacked.

One might argue that the comparison is incongruent. Measuring a sixteenth-century Mughal policy and a fourteenth-century Javanese poem against modern democracies is unjustified, since neither was intended for such a standard. While such a claim holds valid ground, it misses the core argument of the paper. The point isn't that Sulh-i-Kul or Bhinneka Tunggal Ika failed by modern standards, but that they exceeded their limits as institutional measures and poetic metaphors in their attempts to establish pluralism within governance. Their failure stems from a historical pattern: virtue cannot enforce itself, and law outlives its founder.

A brief comparative study allows us to refine the thesis further. The Ottoman Empire faced a similar challenge in ruling a diverse society, yet the response it found was inherently different from Sulh-i-Kul. The Ottomans adopted the Millet System which allowed each officially recognized religious group – Muslims, Orthodox Christians, Armenians and Jews – to govern their own affairs independently. This means that issues related to marriage and inheritance were resolved according to rules and regulations laid down by each specific community's religious body, while the Ottoman state largely assumed the role of a supervising entity (Aviv).

While Sulh-i-Kul was a uniform legal governance policy under a beneficent ruler, the ottoman Millet System delegated legal power among distinct communal groups. Both were pre-modern; however, neither achieved the structural permanence of the later constitutional model, which ensured pluralism through specific legal codes, rather than depending on the benevolence of a ruler or cultural sovereignty. When compared, both Sulh-i-Kul and Bhinneka Tunggal Ika expose their limitations in very similar fashion. The former possessed

characteristics of law, but lacked form – which ultimately perished with its founder. While the latter managed to acquire a form of law, its content remained vulnerable to reinterpretations and adaptation.

E. Conclusion

This paper argues that pluralism will endure only when it shifts from virtue to law. However, the law must still confront its own Aurangzeb. A law can be discontinued; a policy can be dismantled or a clause in a constitution can be reinterpreted to justify a sovereign's greed and wrongdoing. Even the clauses guaranteeing freedom of religion, such as Article 25 of the Indian Constitution or Indonesia's legal protection of Bhinneka Tunggal Ika under Law No.12 of 2011 (Constitution of India; Law No. 12 of 2011), are not unassailable and are ultimately at the mercy of judicial reinterpretation. As highlighted before, pluralistic symbols may be repurposed from declarations of religious diversity to ideological tools of jingoism. The legalisation simply changes the signatory authority, transferring power from a single ruler to a democratic body. Thus, codification does not safeguard any doctrine from abuse; it merely changes the authority that sanctions it.

It is apparent that pluralism is not secured by codification, but that in every new generation, the custodian of codified pluralism must face the same burden that even Akbar and Tantular could not evade: the fact that whatever law a state formalizes, it can also be undone by the state. Nonetheless, despite its disadvantages, legalization still matters as it successfully raises the cost and difficulty of reversal, replacing unpredictability of the supreme rulers with slower and more enduring mechanisms of democracy. Therefore, the ultimate test of diverse democracies may not be their decrees, but the survival of the legal commitments long after initial public enthusiasm disappears

Works Cited

Aghababian, Hana. The Kakawin Sutasoma: A Look at "Bhinneka Tunggal Ika" & Perceptions of the Text's Religious Implications. Independent Study Project (ISP) Collection, SIT Graduate Institute, 2015, digitalcollections.sit.edu/isp_collection/2233/.

Aviv, Efrat. "Millet System in the Ottoman Empire." Oxford Bibliographies in Islamic Studies, Oxford UP, 2016, academic.oup.com/reference/62361/reference-article-abstract/554566607.

Binsasi, Norbertus Antoin, et al. "NKRI and Bhinneka Tunggal Ika: A Constitutional and Futuristic Understanding of Sukarno's Politics." Pena Justisia, 2025, garuda.kemdiktisaintek.go.id/documents/detail/5468201.

Bloembergen, Marieke, and Martijn Eickhoff. "Greater Majapahit: The Makings of a Proto-Indonesian Site across Decolonisation, 1900s–1950s." The Politics of Heritage in Indonesia: A Cultural History, Cambridge UP, 2019, pp. 115–158.

Constitution of India. 1950, arts. 25–28.

Constitution of the Republic of Indonesia. 1945, art. 36A, jdih.dpr.go.id/uud-1945-uu-mk-uu-mk-english.pdf.

Eaton, Richard M. "Temple Desecration and Indo-Muslim States." Journal of Islamic Studies, vol. 11, no. 3, 2000, p. 283.

Government Regulation No. 66 of 1951 concerning the National Emblem [Peraturan Pemerintah Nomor 66 Tahun 1951 tentang Lambang Negara]. Lembaran Negara 1951, No. 111, peraturan.bpk.go.id/Details/77815/pp-no-66-tahun-1951.

Law of the Republic of Indonesia No. 12 of 2011 on the Formation of Laws and Regulations. Art. 6(1)(f).

Mallampalli, Chandra. "Jesuits and the Emperor Akbar, 1580–3." South Asia's Christians: Between Hindu and Muslim, Oxford UP, 2023.

Munir, Bakht. "The Mughal Administration of Justice: An Appraisal." Global Strategic & Security Studies Review, vol. 5, no. 3, 2020, p. 43.

Museum Nasional Indonesia. "Kakawin Sutasoma." 2022, www.museumnasional.or.id/4008/.

Richards, John F. The Mughal Empire. The New Cambridge History of India I.5, Cambridge UP, 1993.

Ricklefs, M. C. A History of Modern Indonesia Since c. 1200. 4th ed., Stanford UP, 2008.

Wolters, O. W. History, Culture, and Region in Southeast Asian Perspectives. Rev. ed., Cornell U Southeast Asia Program Publications, 1999..

Zehra, Qamrush, and Jyotsana Singh. "Slaughter of Animals in Special Reference to Akbar." Indian Journal of Law and Legal Research, vol. 4, no. 2, 2022, p.1

Why Habits Override Conscious Decisions? By Aanya Sodhi

You wake up ready to be productive and finish your work but somehow end up scrolling on your phone for an hour instead. This shared experience can be explained and backed up by research done on the brain to understand the dominant habit system and the gap between intention and action. This is due to the interaction between the two systems of thinking and helps us answer why habits override conscious decisions? On one hand habits are a lot quicker and more efficient than conscious decisions and are strengthened through repetition and reward, while on the other hand, conscious decisions are slower and are an effortful process, which defaults the brain to habitual behaviors.

Habits are learned behaviors that, through repetition, are reinforced and become automatic. They are triggered by context cues, and once fully formed, are insensitive to short-term goal changes and difficult to change. Habit formation is a process that follows the habit loop: cue, routine, and reward. A cue triggers a behavior, which is the routine, and gets reinforced by rewards, increasing the likelihood of it occurring again and then becomes automatic over time. Studies conducted by Charles Duhigg show that 43% of our daily actions happen in the same context, reinforcing our habitual patterns.

A key part in habit formation is dopamine, which plays a crucial role in reinforcement. Dopamine signals in the brain strengthen context-response associations. However, over time dopamine responses decrease because the rewards are expected, shifting the focus from the reward to the cue itself. This explains why individuals may continue a habit even though the reward is reduced or not available. This aligns with research showing that "the brain circuit involved in wanting is much more powerful than the brain circuit involved in satisfaction," (Center for Humane Technology) highlighting how craving and motivation actually outweighs the actual pleasure, therefore reinforcing habitual behavior even more.

When actions continue to persist even though the value of the reward has decreased, it then becomes a habit. There are two systems that control our behavior: System 1 and System 2. The habit system, System 1, is controlled by the Basal Ganglia. It operates automatically and is responsible for enabling fast, efficient behavior that does not require much cognitive effort. It does so by a process called "chunking", which involves breaking down large, complex information into smaller, more manageable parts. It is a bottom-up process as it relies on environmental triggers and on internal feedback. On the other hand, the conscious side is controlled by the Prefrontal Cortex in System 2, which is also known as the goal-directed system. System 2 is a lot slower, requires more cognitive effort and is used for problem-solving scenarios and evaluating outcomes. It uses up a lot of energy to make decisions, which reduces the levels of glucose in the brain. Glucose is the primary fuel for the brain, and this depletion of glucose levels is the reason behind why deep-thinking leaves people feeling tired and fatigued. In fact, in *"Thinking, Fast and Slow",* it states, "The nervous system consumes more glucose than most other parts of the body" (Kahneman, 2011, p. 37), highlighting the energy cost of effortful thinking.

The interaction between these systems can be described with the Dual Process Theory. The brain's primary goal is to conserve energy and so it prefers the automatic system because it is a lot quicker and more efficient. As humans become more stressed, the brain wants to preserve energy and so it defaults on habitual behavior. Although, if needed, the prefrontal cortex can override habits, if it has sufficient motivation and attention.

There are many reasons as to why habits override conscious decisions. Firstly, fatigue plays a huge role in deciding which system dominated behavior. After making many choices, the brain gets tired and goes into "power-save mode" as the prefrontal cortex becomes less effective. So, it starts to rely on automatic processes and habits that are already ingrained into the brain. Secondly, when attention is divided, the brain cannot fully engage and so it increases the reliance on automatic behaviors. An example of multitasking would be watching tv while doing homework, which limits the brain's ability to make conscious decisions. Another instance when habits would take over would be when there is some sort of time pressure because conscious decisions need time to decide and plan, resulting in the brain depending on habits again. Lastly, when individuals are stressed or under cognitive load, they cannot engage in deliberate thinking. For example, students may revert to checking their phone when they are stressed. Habits also can operate in an "open loop", which means that once a behavior has started, it will continue to do so without a second thought, even if it's not the best option anymore. This leads to habit slips, which is when individuals perform actions without meaning to.

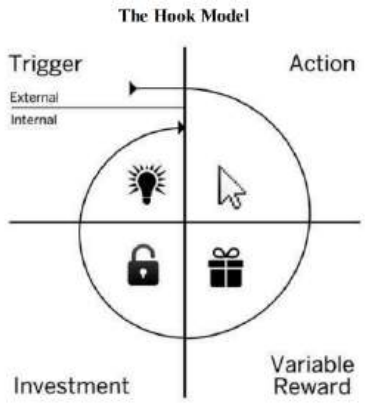
Some real-world examples to better illustrate this are phone usage and nail-biting. They both follow a habit loop. For phone usage, the cue is a notification, checking the phone is the routine, and the reward is the social interaction that you get or the dopamine that is released in your brain as a result. Over time, the cue triggers the response even before your conscious thought can intervene. Similarly, nail-biting is usually triggered by stress or boredom, giving individuals temporary relief in the moment as a reward. Since this behavior has been reinforced several times, the brain defaults to this as it requires little effort. These examples help show how habits can override conscious decisions, even if it may not be the healthiest thing to do.

Businesses have learned to use these psychological concepts to their advantage to increase consumer engagement, profitability, and retention. There are many reasons why businesses want habits. An example is because habits increase customer lifetime value (CLTV) and forces users to continue using products without a second thought. CLTV is the amount of money made from a customer before they die, switch to a competition, or stop using that product. Businesses understand that altering consumer behavior has its challenges and expenses, and so they focus on creating recurring behaviors that are continuously reinforced. This supports the idea that "fostering consumer habits is an effective way to increase the value of a company by driving higher customer lifetime value (CLTV)" (Eyal), since repeated engagement increases a company's revenue. An interesting example of how habit formation is so influential in long-term behavior is the QWERTY keyboard layout. Although the Dvorak keyboard layout, with the vowels in the middle, is proven to be faster and more efficient, most users still use QWERTY because it has been ingrained thoroughly through years of repetition and practice.

Switching to a different system would require relearning how to type, which involves significant cognitive effort and a decrease in performance in the short run. This works as an explanation for one of the biggest reasons why businesses fail: high costs of changing user behavior.

Similar to the habit loop previously mentioned, that relies on reinforcement and repetition, businesses use the Hook Model created by Nir Eyal. According to *Figure 1*, this model has four parts to it: trigger, action, variable reward, and investment. A trigger can be external, like a notification or an alert, or internal, including boredom, loneliness, or stress. The action part covers how companies reduce user effort and increase motivation. Some examples of this would be the infinite scrolling to autoplay feature. They are successful because they require low cognitive effort and rely on system 1 thinking. Variable rewards are unpredictable rewards that increase dopamine activity and are more effective than predictable rewards. Likes, comments, and gambling are examples of variable rewards because they create anticipation, which leads to an increase in cravings and the repeated checking of platforms. This reveals that the brain loves the possibility of a reward, not just the reward. Last but definitely not least, investment is the most important part of the Hook Model. The more effort users invest into a platform, the harder it becomes to leave the platform. Snapchat streaks are a particularly effective example in today's generation because they take advantage of user investment. Streaks lasting hundreds or even thousands of days make missing a single day feel like a waste of all the time and effort they have already invested. As a result, sending a daily snap becomes more of an automatic behavior and less of a conscious decision. Similarly, creating playlists is how music streaming platforms encourage investment. As users spend significant time creating personalized playlists and organizing songs, this makes switching to a competitor much more work and less attractive because they would lose the effort they have already put in.

Figure 1: The Hook Model



Another important concept that was introduced by Nir Eyal was the distinction between painkillers and vitamins. Painkillers solve problems immediately, making it easier to sell as consumers see obvious results. The demand for these types of products is already there as people are willing to consistently pay for them. An example of a painkiller would be the promise that taking a Tylenol will relieve a headache. On the other hand, vitamins' usage is seen as more optional as it provides less urgent benefits, increasing the inconsistency of usage. A famous example of a product in the vitamin category would be Duolingo. This is because it isn't ever urgently needed but it's a good thing to have. Learning Spanish today doesn't solve an urgent problem for most people, rather, the benefits come weeks, months, or years later. Keeping this in mind, every business wants their products to be painkillers because it guarantees consistent usage. In reality, however, many successful products begin as vitamins and use habit-enforcing strategies to become a habit. Through notifications, personalized content, and rewards, products become a bigger part of daily routines until the usage becomes automatic.

After understanding the different strategies that businesses use to design habit-forming systems, they can be seen in multiple real-world applications. Fitbit serves as a clear example of how businesses apply these strategies to transform consumer behavior into automatic routines. Gamification and feedback systems are the two main components that Fitbit uses, in fact, "gamification motivates 89% of users to exercise" (Strivecloud). Streaks, badges, goal-setting systems, step tracking, and other features, help make health decisions into measurable progress. This acts as a reinforcement loop for the brain, strengthening motivation by short-term rewards and indicators of success. In addition, Fitbit also incorporates social reinforcement mechanisms that strengthen accountability. By letting users compare their performance with their friends, join groups, and participate in challenges, it introduces a competitive element. As a result, the behavior transforms from being self-directed to being influenced by external feedback from the outside world. Fitbit demonstrates how successful products are not designed and programmed for the initial attraction, but to become part of users' daily routines. By reinforcing our brain with the small features, rewards, and feedback loop, the platform forces it into a habitual process.

Another key example of businesses using habit-forming strategies is Starbucks. Starbucks came out with its highly personalized mobile app that includes a reward program, as customers can earn stars for purchases and later use them for free drinks and food, challenges, bonus offers, and promotions. Starbucks very efficiently reduces friction, meaning it makes behaviors easier and faster. For example, it offers features like online ordering, digital payment, recent orders, and many more. These features make returning to Starbucks very rewarding and convenient, leading to an increase in its usage and incorporation into their habits. Ultimately, even something as basic as buying a cup of coffee changes from a deliberate decision to a habit, illustrating how companies apply these concepts to increase customer loyalty.

While habit-forming technologies can increase engagement and profitability for businesses, they also have significantly impacted Gen Z. Gen Z is the focus population for my paper because they are the first generation to be completely immersed in technology, social media, and personalized algorithms. The reason behind why they are so susceptible to these

features is because it targets the prefrontal cortex, which develops later than the emotional and reward-processing system in adolescents, making them more vulnerable to persuasive technologies. This is supported by research that using technology during school, “coincides with a developmental stage when attention, self-regulation, and impulse control are still forming” (Hoppe and Elkins). Social media, that we are all surrounded by, is intentionally designed with notifications and visual alerts that act as external triggers to keep reverting users back to the platform. These features activate the brain’s salience network, which makes staying focused a lot harder.

In addition, Gen Z students face immense academic pressure, struggles balancing extracurriculars and future planning. As mentioned earlier, stress weakens the effectiveness of the prefrontal cortex and increases our reliance on habitual behaviors. Students usually turn to social media or gaming to temporarily relieve their stress. These behaviors are reinforced by features such as notifications, likes, and infinite scrolling. After practicing resistance for a long period of time and, “after exerting self-control, decision-making shifts toward simpler and easier processes,” (Weir) which helps explain why students default to behaviors that require low effort. This demonstrates how the same mechanisms used by businesses also influence the decision-making process of the Gen Z population.

In conclusion, habits override conscious decisions because they are built for efficiency, energy conservation and repetition, whereas conscious decision making requires a lot more cognitive effort and is easily susceptible to stress and fatigue. System 1 and 2 interact to shape our behavior and shift the reliance between the systems accordingly. Over time, with reinforcement and repetition, habits become deeply encoded, making them even more resistant to change even if goals shift. The same psychological mechanisms are used by companies to reduce friction, increase customer engagement, and long-term retention, leading to an ultimate increase in customer lifetime value. For Gen Z, the interaction between the two systems is specifically significant because of the developmental stage of the brain and the constant pressure they are under. Understanding how these systems interact helps explain daily behavior and highlighting the vast influence of technology and design on behavior.

Works Cited

Baruch-Feldman, Caren. “Making New Habits and Building Willpower - It’s Not Just for Dieting.” Albert Ellis Institute, albertellis.org/2012/09/making-new-habits-and-building-willpower-its-not-just-for-dieting/.

Blain, Tiara. “Habit Stacking for ADHD.” Verywell Mind, 14 Oct. 2014, www.verywellmind.com/habit-stacking-definition-steps-benefits-for-adhd-6751145.

Cambridge, Mass. “MIT Researcher Sheds Light on Why Habits Are Hard to Make and Break.” MIT News | Massachusetts Institute of Technology, 20 Oct. 1999, news.mit.edu/1999/habits.

Center for Humane Technology. “Social Media and the Brain.” Www.humanetech.com, 17 Aug. 2021, www.humanetech.com/youth/social-media-and-the-brain.

Clinic, Cleveland. “Relapse (Return to Substance Use).” Cleveland Clinic, 7 Apr. 2025, my.clevelandclinic.org/health/articles/relapse-return-to-substance-use.

Eyal, Nir. O v E R 2 0 0 , 0 0 0 c O P I E S S O L D W O R L D W I D E “a Must-Read for Everyone Who Cares about Driving Customer Engagement.” cdn.penguin.co.uk/dam-assets/books/9780241184837/9780241184837-sample.pdf.

Hoppe, Jill, and Elainah Elkins. “Why Is Gen Z Obsessed with Technology?” The Thomas B. Fordham Institute, 2026, fordhaminstitute.org/national/commentary/why-gen-z-obsessed-technology.

“How Fitbit Uses Gamification to Get People Moving.” Strivecloud.io, 2021, www.strivecloud.io/play/fitbit.

Kahneman, Daniel. Thinking, Fast and Slow. 25 Oct. 2011, dn790002.ca.archive.org/0/items/DanielKahnemanThinkingFastAndSlow/Daniel%20Kahneman-Thinking%2C%20Fast%20and%20Slow%20%20.pdf.

Kennerley, Steven W., and Mark E. Walton. “Decision Making and Reward in Frontal Cortex: Complementary Evidence from Neurophysiological and Neuropsychological Studies.” Behavioral Neuroscience, vol. 125, no. 3, June 2011, pp. 297–317, https://doi.org/10.1037/a0023575.

O’Tousa, David, and Nicholas Grahame. “Habit Formation: Implications for Alcoholism Research.” Alcohol, vol. 48, no. 4, June 2014, pp. 327–35, https://doi.org/10.1016/j.alcohol.2014.02.004.

Perez, Sarah. “Starbucks Rolls out a More Personalized Mobile App along with a Revamped Rewards Program | TechCrunch.” TechCrunch, 12 Apr. 2016, techcrunch.com/2016/04/12/starbucks-rolls-out-a-more-personalized-mobile-app-along-with-a-revamped-rewards-program. Accessed 2 June 2026.

Shpancer, Noam. “Futile Persistence: Why Do We Keep Our Bad Habits?” Psychology Today, 26 Aug. 2020, www.psychologytoday.com/us/blog/insight-therapy/202008/futile-persistence-why-do-we-keep-our-bad-habits.

Stanton, James. “What Marketing Directors Can Learn from Starbucks about Building Daily Consumer Habits – BRANDLAB + Ai.” Brandlab.com.au, 30 May 2026, brandlab.com.au/what-marketing-directors-can-learn-from-starbucks-about-building-daily-consumer-habits-2/.

Wallis, Jonathan D. “Orbitofrontal Cortex and Its Contribution to Decision-Making.” Annual Review of Neuroscience, vol. 30, no. 1, July 2007, pp. 31–56, https://doi.org/10.1146/annurev.neuro.30.051606.094334.

Weir, Kirsten. “The Power of Self-Control.” Https://Www.apa.org, Jan. 2012, www.apa.org/monitor/2012/01/self-control.

Wood, Wendy, and Dennis Rünger. “Psychology of Habit.” Annual Review of Psychology, vol. 67, Jan. 2016, pp. 289–314, https://doi.org/10.1146/annurev-psych-122414-033417.

Yin, Henry H. “From Actions to Habits: Neuroadaptations Leading to Dependence.” Alcohol Research & Health, vol. 31, no. 4, 2025, p. 340, pmc.ncbi.nlm.nih.gov/articles/PMC3860473/.

“From Actions to Habits: Neuroadaptations Leading to Dependence.” Alcohol Research & Health, vol. 31, no. 4, 2025, p. 340, pmc.ncbi.nlm.nih.gov/articles/PMC3860473/.

Who Do We Trust? Attachment, Anthropomorphism, and Age in Human–AI Relations
By Ahana Gandhi

Abstract

This study examined age-related differences in attachment styles, anthropomorphism, and trust in artificial intelligence (AI). The primary aim was to examine how emotional bonds and perceptions of AI systems change among adults at different developmental levels. A total of 92 participants aged 16 to 57 completed an online survey assessing attachment patterns, anthropomorphic perceptions, and trust in AI. Results indicated that younger adults demonstrated significantly higher levels of preoccupied (anxious) attachment compared to older groups. Trust in AI differed significantly across age groups, with middle-aged adults reporting the highest trust. Anthropomorphism did not significantly vary by age. Regression analysis showed that attachment styles did not significantly predict anthropomorphism overall, although dismissive attachment showed a modest individual association that should be interpreted cautiously due to the nonsignificant overall model. These findings suggest developmental differences in attachment anxiety and trust in AI, while anthropomorphic tendencies appear relatively stable across adulthood. Given the modest sample size and cross-sectional design, findings should be interpreted cautiously and warrant replication in larger samples.

Keywords artificial intelligence, trust, anthropomorphism, attachment styles, age differences, developmental psychology.

Introduction

Artificial intelligence (AI) is increasingly embedded in everyday life, shaping how individuals work, communicate, and interact with digital systems. As AI technologies become more socially interactive, researchers have begun to examine whether individuals respond to AI using the same psychological mechanisms that govern human relationships. Three constructs are particularly relevant in this context: attachment, trust, and anthropomorphism. Attachment reflects enduring patterns of emotional bonding and expectations of support; trust refers to confidence in the reliability and dependability of an agent; and anthropomorphism involves attributing human-like characteristics to nonhuman entities. Consistent with the Computers-Are-Social-Actors (CASA) paradigm, research demonstrates that individuals often apply social heuristics and interpersonal expectations when interacting with artificial agents (Nass and Moon 85), suggesting that core interpersonal psychological processes may shape how individuals perceive and trust AI systems.

Attachment

Attachment theory provides a framework for understanding how individuals form emotional bonds and regulate interpersonal relationships. Originally developed to explain infant–caregiver relationships, attachment patterns are shaped by early relational experiences and

influence expectations of support, emotional regulation, and trust throughout adulthood (Bartholomew and Horowitz 226). Adult attachment is commonly categorized into four styles: secure, anxious (preoccupied), avoidant-dismissive, and fearful-avoidant. Individuals with secure attachment tend to exhibit comfort with emotional closeness and autonomy, whereas anxious attachment is characterized by heightened concern about abandonment and a strong desire for reassurance. Avoidant-dismissive attachment involves emotional distancing and discomfort with dependence on others, while fearful-avoidant attachment reflects a combination of attachment anxiety and avoidance.

These attachment orientations influence not only interpersonal relationships but also interactions with nonhuman agents. Research suggests that individuals may transfer relational expectations formed in human relationships to artificial systems. For example, individuals with higher attachment anxiety have been found to exhibit lower trust in AI systems, potentially reflecting heightened sensitivity to uncertainty and reliability concerns (Gillath et al. 3). Similarly, attachment insecurity has been associated with stronger emotional engagement and anthropomorphic perceptions toward artificial agents (Heng and Zhang 1778). These findings suggest that attachment processes may play an important role in shaping how individuals perceive and respond to AI.

Trust in AI

Trust is a multifaceted psychological construct that has been defined and operationalized in various ways across disciplines, reflecting its dependence on context, experience, and relational dynamics. Broadly, trust refers to a willingness to accept vulnerability based on positive expectations regarding the intentions or behavior of another agent. Developmental research suggests that trust is shaped early in life through interactions with caregivers, forming internal working models that influence expectations of reliability and support in later relationships. Individuals with consistent and responsive caregiving experiences are more likely to develop secure expectations of others, whereas inconsistent or unreliable caregiving may foster skepticism and reduced willingness to rely on external agents. These foundational trust expectations extend beyond interpersonal relationships and may also shape how individuals evaluate and interact with nonhuman systems, including artificial intelligence, particularly when those systems display socially interactive features.

However, trust in AI differs from interpersonal trust in important ways, as AI systems operate based on programmed algorithms rather than independent intentions. As a result, scholars have argued that trust in AI is often more accurately conceptualized as confidence in system reliability and performance rather than traditional interpersonal trust (Ryan 2751; Körber 15). Recent human-robot interaction research also shows that trust in AI depends not only on system performance but on human factors, including prior experiences and social expectations. Measures of trust in automation emphasize perceived competence, predictability, and responsibility, which align closely with the interpersonal trust mechanisms shaped by attachment history.

The dependability and reliability of AI systems remain important research concerns. As artificial intelligence becomes more involved in daily life, there have been more questions about the dependability and reliability of AI. These systems are trained to give quick responses, but the validity of the response is contingent on the data it is trained on (Hong et al. 60). The quality of the data can vary, and incomplete, inaccurate, and biased datasets can easily impact AI responses and negatively affect dependability. It is important to note that artificial intelligence does not possess human-like understanding but instead generates outputs based on learned statistical patterns derived from training data, which can result in inaccurate or misleading outputs under certain conditions. Another aspect that affects the dependability of AI is algorithmic transparency. Many AI systems operate as so-called "black boxes," meaning their internal decision processes are not fully transparent to users or developers, which makes it difficult to assess the credibility of these tools. Lack of accuracy and general ambiguity about the sources behind the information that AI provides can lead to less trust in AI, which would in turn lead to less dependability or reliability (Ryan 2751). Trust calibration research further finds that users adjust their trust when AI errors accumulate, when explanations are missing, or when transparency decreases (de Visser et al. 1415), highlighting that dependability is a dynamic perception rather than a fixed system property.

Anthropomorphism and Attachment to AI

Anthropomorphism is the tendency to attribute human-like characteristics to nonhuman entities, and it plays a significant role in the way people interact with artificial intelligence (Ryan 2751). As AI systems become more advanced, they are designed to mimic human behaviors, emotions, and expressions, making them seem more relatable and trustworthy. This has led to a growing phenomenon where people develop emotional attachments to AI-powered entities, such as chatbots, virtual assistants, and social robots. Research consistently shows that anthropomorphic cues increase perceived warmth, competence, and trust in AI systems (Deng and Yan 4). Because AI does not possess genuine emotional or cognitive experience, anthropomorphism toward AI reflects users' psychological perceptions rather than intrinsic properties of the system.

Psychological research suggests that attachment to AI may stem from similar mechanisms as human attachment. People who struggle with human relationships or who experience social isolation are more likely to form strong emotional bonds with AI (Zimmerman et al. 1559). This is particularly evident in human-robot interactions, where users perceive AI as a comforting presence. While these attachments can provide companionship, they also raise ethical and psychological concerns: over-dependence on AI companionship could impede genuine social development and exacerbate avoidant behaviors. Furthermore, because AI does not possess genuine emotional experience, perceived emotional reciprocity reflects users' psychological interpretation rather than actual emotional exchange. Understanding the implications of anthropomorphism and AI attachment is crucial in developing ethical AI systems that promote healthy interactions without fostering emotional dependence.

This research investigates the dynamics of emotional attachment, trustworthiness, dependability, and reliability within the context of artificial intelligence systems, examining both the underlying factors and possible risks related to these psychological and technical aspects. With AI's continuous advancement and growing presence across multiple domains of everyday existence, the nature of human-AI interactions has become progressively more complex. People exhibiting insecure attachment patterns, especially those demonstrating anxious or avoidant characteristics, may experience difficulties in interpersonal relationships and consequently form emotional connections with AI-powered systems, including digital assistants, conversational agents, or robotic companions with human-like features. Although such connections may offer a form of social support, they could simultaneously promote isolation from society and impede the growth of authentic emotional capacities in face-to-face interactions.

Methodology

Aim. The present study aims to investigate how attachment styles, anthropomorphism, and trust in AI differ across age groups, and to explore the interrelationships among these psychological constructs in shaping human–AI interaction.

Hypotheses. H1–H4: Secure, fearful, preoccupied, and dismissive attachment styles will differ significantly across age groups. H5: Anthropomorphism will differ significantly across age groups. H6: Trust in AI will differ significantly across age groups. H7: Attachment, anthropomorphism, and trust in AI will be correlated. H8: Attachment styles will significantly predict anthropomorphism toward AI.

Supplementary Hypotheses. In addition to the primary inferential hypotheses, supplementary hypotheses guided the interpretation of the descriptive figures presented in the Results section. H9: Respondents will report predominantly moderate-to-high levels of positive emotion (excitement, optimism, relaxation) toward AI. H10: Respondents will report predominantly low-to-moderate levels of negative emotion (worry, fear, outrage) toward AI. H11: Respondents will most frequently report a neutral level of willingness to trust AI. H12: Respondents will report moderate-to-high levels of acceptance of AI use. H13: Respondents will report moderate-to-high levels of perceived familiarity with AI. H14: Respondents will most frequently endorse neutral or mixed beliefs regarding whether AI will create more jobs than it eliminates. H15: One AI platform will emerge as the most frequently used tool among respondents.

Participants. A total of 92 individuals participated in the study. Data were collected from May to July 2025 using an online survey distributed through a snowball sampling strategy. Inclusion criteria required participants to be 16 years or older and proficient in English. Although an a priori power analysis indicated that a minimum of 159 participants would be required to reliably detect medium effect sizes, the final sample consisted of 92 participants who completed the survey in full and met inclusion criteria. Participants were between the ages of 16 and 57, with 30.5% between 16 and 25, 31.5% between 26 and 41, and 38.1% between 42 and 57 — age brackets chosen to reflect recognized developmental and psychosocial differences among early,

middle, and late adulthood. The sample consisted of 59.4% females and 40.6% males; 22.9% had a high school diploma, 27.1% an undergraduate degree, 43.8% a postgraduate degree, and 6.3% a doctorate. 70.8% were employed, 24% were students, 3.1% were not working, and 2.1% were skilled professionals.

Instruments. Trust-related perceptions of AI were assessed using a 9-item scale adapted from the KPMG Trust in Artificial Intelligence Survey, covering data privacy, transparency, accountability, fairness, and human oversight on a 5-point Likert scale. The adapted scale demonstrated excellent internal consistency (α = .94). Anthropomorphism was measured with a self-constructed 10-item scale evaluating cognitive and affective anthropomorphism (e.g., "AI seems to have a mind of its own"), based on the framework of Waytz, Cacioppo, and Epley (219). This scale showed acceptable internal consistency (α = .79). Attachment was measured using the 20-item Attachment Styles Questionnaire (Van Oudenhoven et al.), based on Bartholomew and Horowitz's four-category model, yielding secure (α = .84), fearful (α = .85), preoccupied (α = .67), and dismissive (α = .60) subscales.

Table A. Internal Consistency Reliability of the Anthropomorphism Scale (10 items, overall α = .79)

Item	Corrected item-total r	α if item deleted
ANT_1	0.28	0.79
ANT_2	0.57	0.76
ANT_3	0.43	0.77
ANT_4	0.25	0.79
ANT_5	0.38	0.78
ANT_6	0.56	0.75
ANT_7	0.47	0.77
ANT_8	0.59	0.75
ANT_9	0.53	0.76
ANT_10	0.52	0.76

Corrected item-total correlations ranged from .25 to .59, with most items exceeding the recommended .30 threshold. Items ANT_1 and ANT_4 showed relatively lower item-total correlations, but their removal did not improve overall reliability, supporting retention of all ten items.

Table B. Internal Consistency Reliability of the Trust Scale (9 items, overall α = .94)

Item	Corrected item-total r	α if item deleted
TRUST_1	0.76	0.94
TRUST_2	0.80	0.943
TRUST_3	0.81	0.94
TRUST_4	0.86	0.93
TRUST_5	0.84	0.93
TRUST_6	0.78	0.94
TRUST_7	0.76	0.94
TRUST_8	0.69	0.94
TRUST_9	0.77	0.94

Corrected item-total correlations ranged from .69 to .86, well above the recommended minimum, and removal of any single item would not have meaningfully improved reliability, confirming the Trust Scale's strong internal coherence.

Procedure and Analysis. Participants completed an online Google Forms questionnaire after providing informed consent, which took approximately 10–12 minutes. Data were analyzed using DataTab statistical software. One-way ANOVAs examined age-group differences in attachment, anthropomorphism, and trust, with Tukey's HSD post hoc tests for significant effects. Pearson correlation analysis examined relationships among constructs, and multiple linear regression tested whether attachment styles predicted anthropomorphism. Assumptions of normality (Shapiro–Wilk, Kolmogorov–Smirnov, Anderson–Darling) and homogeneity of variance (Levene's test) were satisfactorily met for all analyses.

Ethical Considerations. The study followed established ethical guidelines for human subjects research. Before taking part, participants gave informed consent electronically. Participants were assured of anonymity and confidentiality and were informed that they could withdraw at any time without consequence. Survey items were designed to minimize psychological risk, and participants were provided with contact information for further inquiries or concerns. The dataset and analysis scripts used to generate the reported results are available from the corresponding author upon reasonable request, in order to protect participant confidentiality given the nature of the survey items.

Results

Descriptive Findings. Figures 1 through 11 summarize respondents' emotional reactions to AI, trust, acceptance, familiarity, beliefs about employment impact, and platform usage.

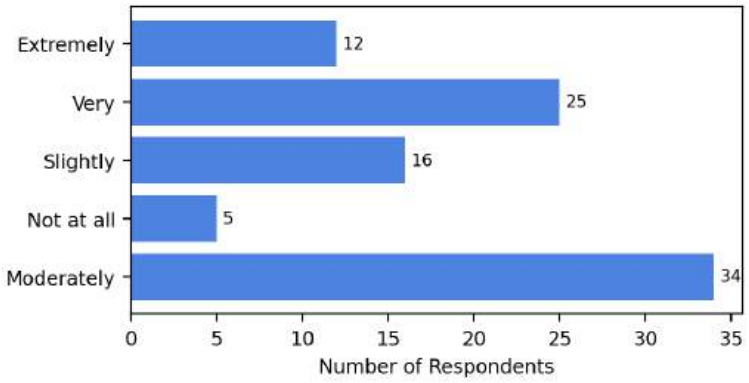


Fig. 1. The extent to which respondents feel excited about AI.

The majority of respondents (n = 34) reported feeling moderately excited about AI, while n = 25 reported feeling very excited. Only a small number (n = 5) reported not feeling excited at all.

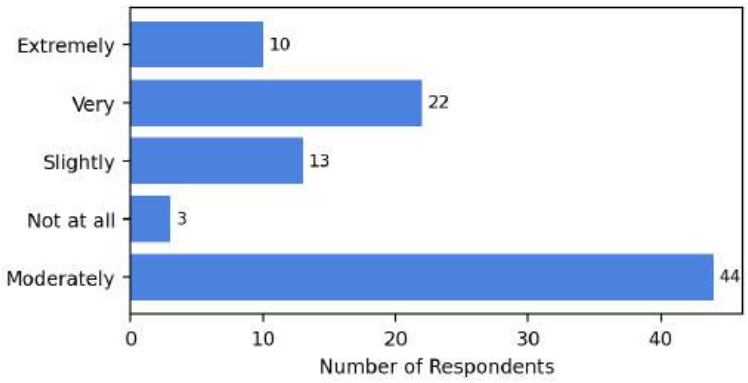


Fig. 2. The extent to which respondents feel optimistic about AI.

Most participants (n = 44) reported feeling moderately optimistic about AI, making it the most common response, with a substantial number (n = 22) feeling very optimistic. Only a small minority (n = 3) reported no optimism at all.

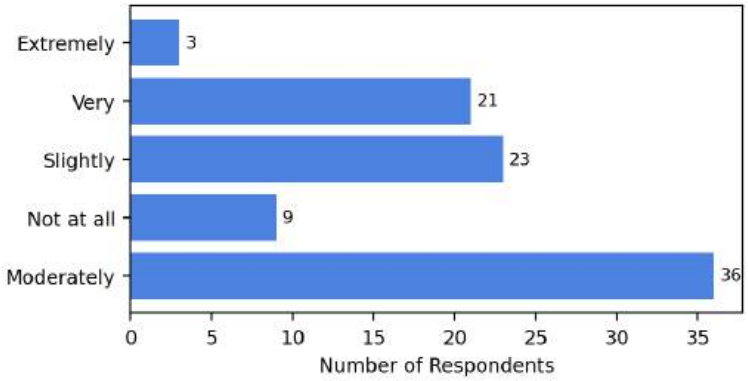


Fig. 3. The extent to which respondents feel relaxed about AI.

Most participants (n = 36) reported feeling moderately relaxed when thinking about AI, followed by those who felt slightly relaxed (n = 23). A smaller portion (n = 21) felt very relaxed, while very few (n = 3) felt extremely relaxed.

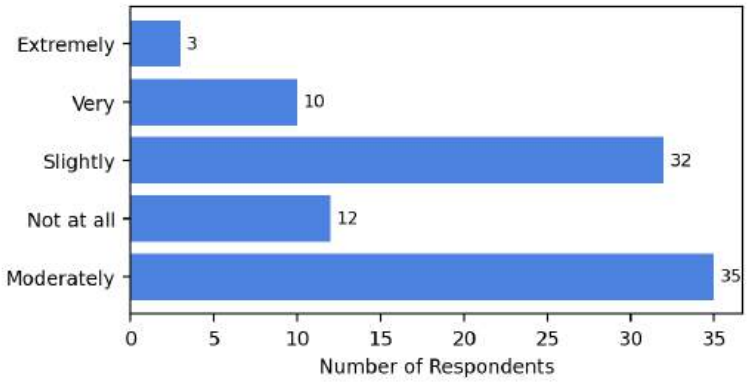


Fig. 4. The extent to which respondents feel worried about AI.

The two most common responses were moderately worried (n = 35) and slightly worried (n = 32), suggesting mild-to-moderate concern was relatively widespread. Only a minimal number (n = 3) reported feeling extremely worried.

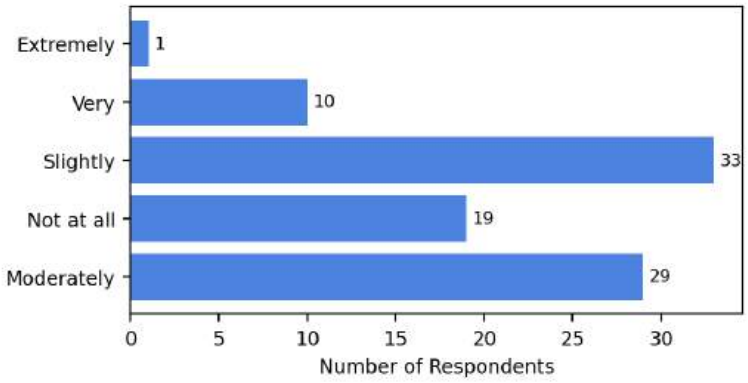


Fig. 5. The extent to which respondents feel fearful about AI.

The most common response was slightly fearful (n = 33), followed by moderately fearful (n = 29). A notable portion (n = 19) reported feeling not at all fearful, and only one respondent reported feeling extremely fearful.

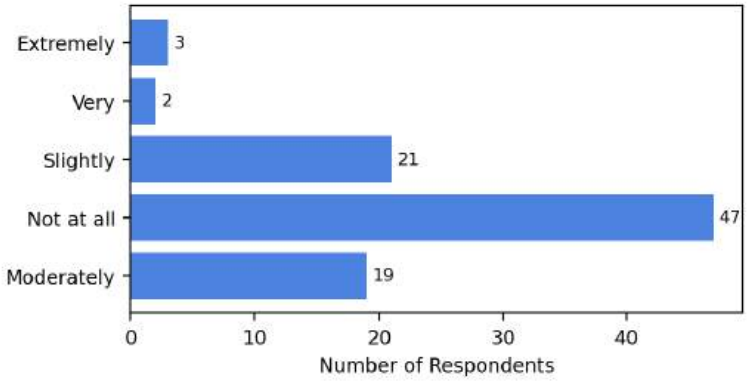


Fig. 6. The extent to which respondents feel outraged about AI.

Feelings of outrage toward AI were generally low: nearly half of respondents (n = 47) reported feeling not at all outraged, while very few (n = 3) reported feeling extremely outraged.

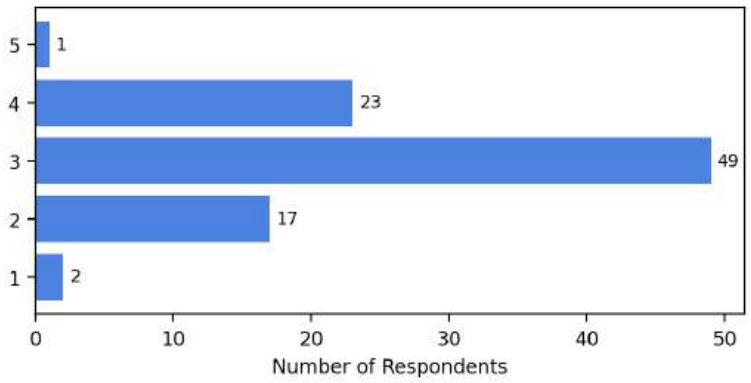


Fig. 7. Respondents' willingness to trust AI (1 = very low trust, 5 = very very high trust).

The majority of respondents (n = 49) selected the neutral midpoint of the trust scale. A notable proportion (n = 23) reported a relatively higher level of trust, while only small groups indicated minimal trust (n = 2) or very high trust (n = 1).

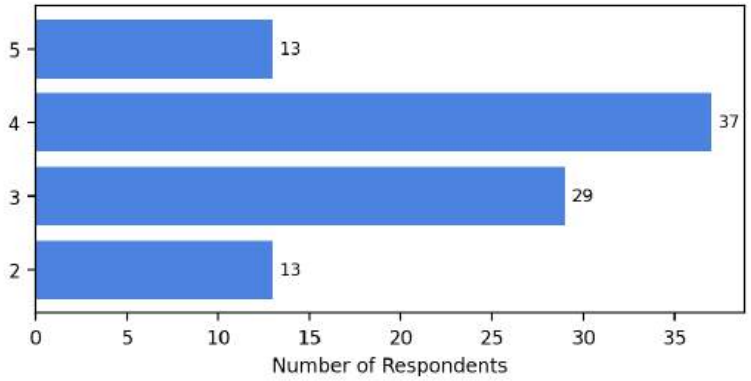


Fig. 8. The extent to which respondents accept the use of AI.

The most common level of acceptance was 4 out of 5 (n = 37), followed by the neutral midpoint (n = 29). No respondents selected the lowest level, suggesting outright rejection of AI use was absent from the sample.

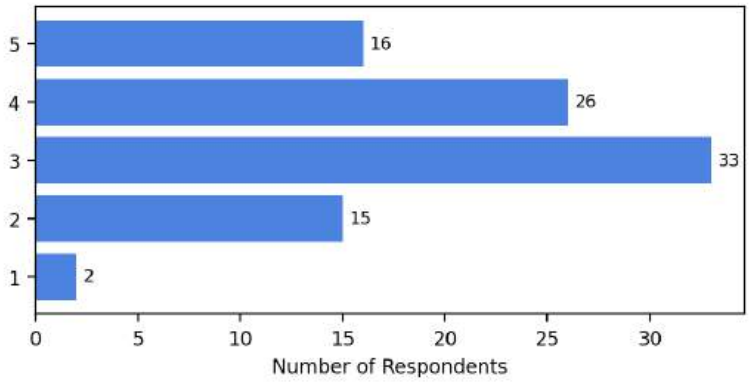


Fig. 9. Respondents' self-reported familiarity with AI.

The most common familiarity rating was the midpoint (n = 33), followed by respondents who rated their familiarity as relatively high (n = 26) or very high (n = 16), suggesting a generally well-informed respondent group.

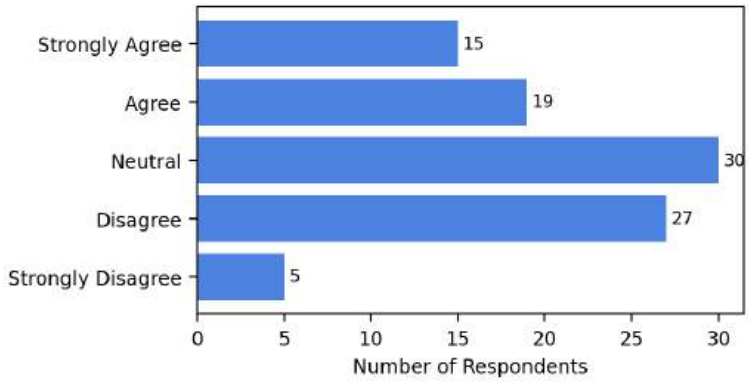


Fig. 10. The extent to which respondents believe AI will create more jobs than it will eliminate.

The largest proportion of participants selected Neutral (n = 30), indicating uncertainty about AI's net impact on employment, while a substantial number (n = 27) disagreed that AI would create more jobs than it eliminates.

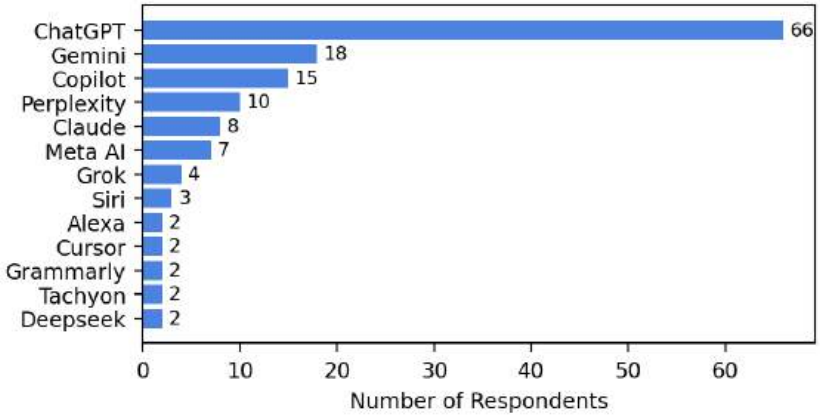


Fig. 11. AI platforms commonly used by respondents.

Since participants could select more than one platform, totals exceed the number of respondents. ChatGPT was used by a substantially larger proportion of respondents (n = 66) than any other platform, followed by Gemini (n = 18) and Copilot (n = 15); Perplexity, Claude, and Meta AI saw moderate use, while several other tools were each used by only a handful of respondents.

Taken together, these descriptive patterns suggest that respondents generally viewed AI with more positive than negative emotion, held a cautiously neutral-to-positive stance on trusting and accepting AI, considered themselves reasonably familiar with the technology, and were most likely to use general-purpose conversational tools such as ChatGPT. These patterns supported H9 through H15: positive emotion was predominantly moderate-to-high, negative emotion

predominantly low-to-moderate, trust modal at neutral, acceptance and familiarity moderate-to-high, beliefs about job creation predominantly neutral, and one platform (ChatGPT) clearly dominated usage.

Age Differences in Attachment Styles

Table 1. ANOVA Results for Age Groups and Attachment Styles

Attachment style	16–25 M (SD)	26–41 M (SD)	42–57 M (SD)	F(2,89)	p	η^2
Secure	27.04 (3.88)	26.48 (4.90)	26.91 (4.33)	0.15	.881	0.00
Fearful	11.57 (3.02)	11.86 (3.93)	10.37 (3.37)	1.69	.19	0.04
Preoccupied	19.54 (5.83)	14.69 (3.85)	16.83 (3.89)	8.82	<.001	0.17
Dismissive	11.14 (2.24)	11.62 (2.06)	11.09 (1.99)	0.60	.553	0.01

Secure, fearful, and dismissive attachment styles did not differ significantly across age groups. Preoccupied (anxious) attachment differed significantly by age, $F(2, 89) = 8.82$, $p < .001$, $\eta^2 = 0.17$, with the youngest group (16–25) scoring markedly higher than the oldest group (42–57). Tukey's HSD post hoc comparisons (Table 2) confirmed that the 16–25 group scored significantly higher than the 42–57 group; the difference between the 16–25 and 26–41 groups approached significance.

Table 2. Tukey's HSD Post Hoc Comparisons for Preoccupied Attachment

Comparison	Mean diff.	p	95% CI
16–25 vs. 42–57	4.85	<.001	[2.10, 7.60]
16–25 vs. 26–41	2.71	.069	[-0.17, 5.58]
42–57 vs. 26–41	2.14	.153	[-0.58, 4.87]

Age Differences in Anthropomorphism and Trust in AI

Table 3. ANOVA Results for Age Groups, Anthropomorphism, and Trust in AI

Variable	16–25 M (SD)	26–41 M (SD)	42–57 M (SD)	F(2,89)	p	η^2
Anthropomorphism	26.86 (7.51)	26.83 (5.96)	25.46 (6.33)	0.48	.621	0.01
Trust in AI	36.68 (8.13)	40.03 (5.67)	34.06 (8.13)	5.11	.008	0.10

Anthropomorphism did not differ significantly across age groups. Trust in AI differed significantly by age, $F(2, 89) = 5.11$, $p = .008$, $\eta^2 = 0.10$, with the 26–41 group reporting the highest trust and the 42–57 group the lowest.

Table 4. Tukey's HSD Post Hoc Comparisons for Trust in AI

Comparison	Mean diff.	p	95% CI
16–25 vs. 42–57	2.62	.351	[-1.88, 7.12]
16–25 vs. 26–41	3.36	.21	[-1.35, 8.06]
42–57 vs. 26–41	5.98	.005	[1.52, 10.43]

Post hoc comparisons showed that the 26–41 group scored significantly higher on trust in AI than the 42–57 group (p = .005). No significant differences emerged between the 16–25 group and either of the other two age groups.

Correlations Among Attachment, Anthropomorphism, and Trust in AI

Table 5. Pearson Correlations Among Attachment Subscales, Anthropomorphism, and Trust in AI

	Secure	Fearful	Preoccup.	Dismiss.	Anthro.	AI Trust
Secure	—	-.52***	-.22*	.03	-.03	.12
Fearful	-.52***	—	.46***	.00	.09	.00
Preoccupied	-.22*	.46***	—	-.22*	.15	-.03
Dismissive	.03	.00	-.22*	—	.19	.07
Anthropomorphism	-.03	.09	.15	.19	—	.17
AI Trust	.12	.00	-.03	.07	.17	—

**p < .05. ***p < .001.*

Secure attachment was significantly negatively correlated with fearful attachment (r = -.52, p < .001) and preoccupied attachment (r = -.22, p = .034). Fearful and preoccupied attachment were significantly positively correlated (r = .46, p < .001), and preoccupied and dismissive attachment were significantly negatively correlated (r = -.22, p = .038). No attachment subscale was significantly correlated with anthropomorphism or trust in AI, and anthropomorphism was not significantly correlated with trust in AI (r = .17, p = .097).

Regression Predicting Anthropomorphism from Attachment Styles

Table 6. Multiple Linear Regression Predicting Anthropomorphism

Predictor	B	SE B	β	t	p
Constant	13.27	7.75	—	1.71	.091
Secure attachment	0.01	0.18	.01	0.05	.958
Fearful attachment	0.01	0.25	.00	0.03	.975
Preoccupied attachment	0.26	0.16	.20	1.65	.102
Dismissive attachment	0.73	0.34	.23	2.19	.031

Note. N = 92. Model fit: $R^2 = .07$, adjusted $R^2 = .03$, $F(4, 87) = 1.74$, $p = .148$.

The overall regression model predicting anthropomorphism from the four attachment subscales was not statistically significant, $F(4, 87) = 1.74$, $p = .148$, $R^2 = .07$. Dismissive attachment was the only individual predictor to reach significance ($B = 0.73$, $\beta = .23$, $p = .031$); because the overall model was nonsignificant, this individual association should be treated as exploratory and requires replication.

Discussion

This study examined age-related differences in attachment, anthropomorphism, and trust in AI, as well as the interrelationships among these constructs. Overall, the findings partially supported the proposed hypotheses. Consistent with H3, significant age-group differences emerged for preoccupied (anxious) attachment, with younger adults (16–25) reporting substantially higher levels of attachment anxiety than middle-aged and older adults, suggesting a developmental decline in attachment-related insecurity. H6 was also supported: trust in AI differed significantly across age groups, with adults aged 26–41 reporting higher levels of trust than older adults, indicating that trust in AI may peak during early-to-mid adulthood. In contrast, H1, H2, and H4 were not supported, as secure, fearful, and dismissive attachment styles did not significantly vary across age groups. Likewise, H5, which predicted age differences in anthropomorphism, was rejected; participants showed similar tendencies to attribute human-like qualities to AI regardless of age. Correlational analyses partially supported H7 by revealing significant associations among several attachment dimensions, although not all predicted relationships with anthropomorphism and AI trust were observed. H8 was also only partially supported: while attachment styles as a set did not significantly predict anthropomorphism, dismissive attachment emerged as a unique significant predictor.

All three age groups had comparable levels of secure attachment as well as dismissive and fearful avoidant attachment. The mean differences were very small, and the ANOVA results were nonsignificant, indicating that being in one's late teens versus one's forties or fifties did not substantially influence how securely or avoidantly attached a person was. This outcome is consistent with research suggesting that adult attachment orientations remain relatively stable across adulthood: prior work has found no significant age differences in secure or avoidant attachment between much younger and much older adult samples, with older adults scoring as high on security and as low on avoidance as younger adults (Segal et al. 125). It is worth noting that the oldest group in this study (42–57) may not have been old enough to capture certain later-life changes in attachment; some longitudinal research suggests a possible uptick in dismissing-avoidant attachment at more advanced ages, which this study's age cap would have missed.

In contrast to the stable patterns for security and avoidance, attachment anxiety showed a pronounced age-related decline consistent with H3. Younger adults had markedly higher preoccupied attachment scores than both middle-aged and older adults, mirroring longitudinal findings that attachment anxiety decreases steadily from young adulthood through middle age (Chopik et al., "Longitudinal Changes" 605). Several factors might explain this developmental trajectory: as people age and accumulate relationship experience, they often develop greater emotional security, self-regulation skills, and confidence in their social bonds, which can reduce fears of abandonment that underlie anxious attachment. Socioemotional selectivity theory also offers insight — as adults age, they tend to prioritize close, positive relationships and downsize their social networks to those who provide the most emotional support, which could naturally contribute to lower attachment anxiety (Charles and Carstensen 390).

Hypothesis 5, anticipating age differences in anthropomorphism, was not supported: anthropomorphism scores were very similar across all three age brackets, and the effect size was negligible. This null finding adds to a limited and inconsistent literature on age and anthropomorphism. Some studies have found that older adults may anthropomorphize robots more than younger adults under certain conditions, potentially reflecting a novelty effect among those with less lifetime exposure to advanced technology (Morillo-Mendez et al. 1075); however, the present sample's oldest participants (in their 40s and 50s) were likely already familiar with modern AI tools, which may have blunted any novelty-based anthropomorphism and resulted in more uniform perceptions across ages.

Lastly, age had a notable effect on trust in AI, supporting H6. The highest trust was reported not by the youngest cohort but by middle-aged adults (26–41), slightly higher than the 16–25 group, while the oldest group (42–57) was the most skeptical. This pattern aligns with broader generational trends: surveys have found that younger consumers are considerably more likely to trust AI than older consumers, and that greater digital exposure and familiarity among younger and middle-aged adults tend to produce more benevolent views of AI, while older adults' privacy concerns and lower digital trust temper their willingness to adopt it (Kubovics 104).

Correlational results partially supported H7: several attachment subscales were meaningfully related to one another, but none were significantly related to anthropomorphism or trust in AI, and anthropomorphism and trust were not significantly related to each other. The regression model testing H8 was not significant overall, though dismissive attachment showed an exploratory individual association with anthropomorphism, consistent with theories that emotionally distant individuals may substitute human-like engagement with predictable, low-risk artificial agents (Heng and Zhang 1780).

Conclusion

The present study explored how attachment styles, anthropomorphism, and trust in AI vary across different adult age groups. Two statistically significant differences emerged: younger adults (16–25) reported significantly higher levels of preoccupied attachment (attachment anxiety), and middle-aged adults (26–41) reported significantly higher levels of trust in AI compared to both younger and older groups. These patterns are consistent with prior psychological and sociotechnical literature, which suggests that attachment anxiety typically declines with age, while trust in technology tends to peak during stages of life when individuals have both exposure and professional engagement with emerging tools like AI. While other attachment styles (secure, dismissive, fearful) and anthropomorphism scores did not significantly differ across age groups, this too is in line with research showing that these variables tend to remain relatively stable throughout adulthood. Together, these results highlight developmental and generational differences in how people emotionally relate to both other humans and intelligent machines, with implications for mental health professionals, human-computer interaction designers, and developers working on AI trustworthiness.

Limitations

Small Sample Size and Sampling Method. With 92 participants, this study's sample size limits the ability to draw broadly generalizable conclusions. Because each age group consisted of a relatively small number of participants, statistical power was reduced, particularly for detecting smaller effects — consistent with the a priori power analysis indicating that a larger sample would be needed to reliably detect medium-sized effects. Snowball sampling also reduces representativeness, as recruitment may be biased toward individuals within shared social networks.

Age Range Restriction. Participants were no older than 57, so this study cannot speak to how attachment and AI trust may look among older adults, who tend to trust AI less than younger participants and may show different attachment patterns in late life.

Self-Report Bias. All measures relied on self-report questionnaires, which are vulnerable to social desirability bias, recall inaccuracies, and limited introspective access to one's own internal states.

Cross-Sectional Design. Because the study used a cross-sectional design, it cannot capture how individual participants' attachment styles or trust in AI might change over time; longitudinal designs would be needed to document such temporal variation.

Lack of Standardized AI-Specific Measures. No widely validated, standardized scales exist for measuring trust or anthropomorphism specifically toward modern conversational AI, so this study relied on adapted measures. While internal consistency was acceptable in this sample, the lack of extensive prior validation may limit comparability with other studies.

Lack of Covariate Controls. This study did not control for potential covariates such as gender, education level, or prior familiarity with AI, which may influence both trust and anthropomorphic perceptions.

Future Research and Recommendations

Future work should pursue longitudinal designs to track how attachment styles, AI trust, and anthropomorphic perceptions evolve over time, and cross-cultural comparisons to test whether these age-related patterns hold across different technological and cultural contexts. Expanding the age range to include adults over 60 would help capture late-life shifts in trust, attachment, and digital literacy. There is also a clear need for empirically validated, AI-specific instruments to measure trust and anthropomorphism, which would improve measurement reliability and enable comparison across studies. Practically, these findings suggest value in designing age-sensitive AI interfaces — for example, emphasizing transparency and reliability cues for older adults and personalization features for younger users — and in implementing AI literacy initiatives that could reduce unwarranted skepticism or overtrust, particularly among older adults with less familiarity with AI systems.

Appendix A: Anthropomorphism Scale

All items rated on a 5-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree).

1. AI seems to have a mind of its own.
2. AI can understand what I am thinking.
3. I believe the AI makes decisions based on its own intentions.
4. I feel that AI has preferences or goals.
5. I would feel bad if I were rude to AI.
6. I talk to AI as if it were a person.
7. I give AI a name or personality.
8. When I interact with AI, I feel like I'm talking to someone, not something.
9. AI deserves to be treated with respect.
10. I would feel guilty if I ignored or disrespected AI.

Appendix B: Trust in Artificial Intelligence Scale

Trust in artificial intelligence was assessed using items adapted from the KPMG Trust in Artificial Intelligence Survey. Participants responded to the prompt: "How important are the following factors for you to trust artificial intelligence systems?" Responses were recorded on a 5-point Likert scale (1 = Not at all important, 2 = Slightly important, 3 = Moderately important, 4 = Highly important, 5 = Extremely important).

1. Data privacy and security
2. Technical robustness and safety
3. Transparency and explainability
4. Risk and impact mitigation
5. Accountability and contestability
6. Human agency and oversight
7. Fairness, inclusion, and non-discrimination
8. AI literacy
9. Ethical code of conduct and governance

Works Cited

Bartholomew, Kim, and Leonard M. Horowitz. "Attachment Styles Among Young Adults: A Test of a Four-Category Model." Journal of Personality and Social Psychology, vol. 61, no. 2, 1991, pp. 226–244.

Charles, Susan T., and Laura L. Carstensen. "Social and Emotional Aging." Annual Review of Psychology, vol. 61, 2010, pp. 383–409.

Chopik, William J., et al. "From the Cradle to the Grave: Age Differences in Attachment from Early Adulthood to Old Age." Journal of Personality, vol. 81, no. 2, 2013, pp. 171–183.

Chopik, William J., et al. "Longitudinal Changes in Attachment Orientation over a 59-Year Period." Journal of Personality and Social Psychology, vol. 116, no. 4, 2019, pp. 598–611.

de Visser, Ewart J., et al. "From 'Automation' to 'Autonomy': The Importance of Trust Repair in Human–Machine Interaction." Ergonomics, vol. 61, no. 10, 2018, pp. 1409–1427.

Deng, Zhihao, and Jing Yan. "The Effect of Perceived Warmth, Competence, and Social Presence of AI-Driven Chatbots on Consumers' Engagement and Satisfaction." SAGE Open, vol. 15, no. 3, 2025.

Gillath, Omri, et al. "Attachment and Trust in Artificial Intelligence." Computers in Human Behavior, vol. 115, 2020.

Heng, Sok, and Zhi Zhang. "Attachment Anxiety and Problematic Use of Conversational Artificial Intelligence." Psychology Research and Behavior Management, vol. 18, 2025, pp. 1775–1785.

Hong, Yili, et al. "Statistical Perspectives on Reliability of Artificial Intelligence Systems." Quality Engineering, vol. 35, no. 1, 2022, pp. 56–78.

Körber, Moritz. "Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation." Advances in Intelligent Systems and Computing, 2018, pp. 13–30.

Kubovics, Melinda. "The Impact of Age Groups' Attitudes Towards Artificial Intelligence and Data Protection." International Journal of Business and Economics Research, vol. 14, no. 3, 2025, pp. 99–108.

Morillo-Mendez, Lucas, et al. "Age-Related Differences in the Perception of Robotic Referential Gaze in Human-Robot Interaction." International Journal of Social Robotics, vol. 16, no. 6, 2022, pp. 1069–1081.

Nass, Clifford, and Youngme Moon. "Machines and Mindlessness: Social Responses to Computers." Journal of Social Issues, vol. 56, no. 1, 2000, pp. 81–103.

Ryan, Mark. "In AI We Trust: Ethics, Artificial Intelligence, and Reliability." Science and Engineering Ethics, vol. 26, no. 5, 2020, pp. 2749–2767.

Segal, Daniel L., et al. "Age Differences in Attachment Orientations Among Younger and Older Adults." International Journal of Aging and Human Development, vol. 69, no. 2, 2009, pp. 119–132.

Van Oudenhoven, Jan Pieter, et al. "Attachment Styles Questionnaire." Personality and Individual Differences, 2003.

Waytz, Adam, John Cacioppo, and Nicholas Epley. "Who Sees Human?" Perspectives on Psychological Science, vol. 5, no. 3, 2010, pp. 219–232.

Zimmerman, Annette, et al. "Human/AI Relationships: Challenges, Downsides, and Impacts on Human/Human Relationships." AI and Ethics, vol. 4, no. 4, 2023, pp. 1555–1567.

American Youth Desensitization to Political Violence By Leighton Westerberg

Abstract

With the increasing accessibility of political violence on social media, especially following the death of Charlie Kirk, the researcher became interested in the effects of political violence content. This study sought to determine the impact exposure to political violence on social media has on American high school students' emotional desensitization. Using studies based in other countries and different age groups as inspiration, the researcher utilized a mixed-methods approach to collect both quantitative and qualitative data. Students in high schools local to Calabasas, California, were encouraged to answer a Google Forms survey. The researcher also conducted interviews with child psychologists. All data collected was analyzed against three themes: lack of empathy, emotional reaction, and political interest/involvement. The results showed that underclassmen, when compared to upperclassmen, had greater emotional reactions to politically violent content while also reporting that exposure to such content normalized political violence. Implications of this study include the need for digital literacy education programs to encourage healthy boundary-setting and protection from excessive media consumption. It also suggests the need for emotional counseling resources for students to work through and understand the media they consume online, especially that of political violence content.

Political violence desensitization is defined as the "diminished emotional responsiveness with repeated encounters with political violence" (Mrug). With the rise of social media and the digital age, emotional desensitization to political violence is a problem now more than ever. This study sought to analyze the extent to which exposure to political violence on social media affects emotional desensitization among American high school students.

Although social media has been credited for increasing political engagement among youth users, not all political engagement is positive or even elicits a reaction. There are limited studies that have been conducted to examine the relationship between political violence exposure on social media and emotional desensitization, especially among American high school students. According to researcher Desmond Upton Patton, despite limited information regarding this relationship, there is a definite negative connection between social media and emotional desensitization (Patton). Most studies are conducted in other countries using middle school or university students. However, this is a problem for American high school students, too.

With the easy accessibility and frequency of social media posts containing information, graphic content, and public opinions about an act of political violence, students may become desensitized to such content. While official news accounts are required to spread information ethically, accurately, and safely, personal accounts do not have that same responsibility. The lack of names or contact information attached to an account makes posting such content easier, as there are no real-life consequences. These accounts can post graphic content or spread biased opinions as facts without an issue.

Comment sections are also severely underregulated. Many users avoid platform guidelines and express controversial or untruthful thoughts by simply substituting symbols or numbers for letters (i.e. d!e or murd3r). The inability to be punished in real life for the messages users share online further encourages the use of comment sections to express their thoughts.

Social media spreads information without boundaries and can expose youth to politically violent content. Algorithmic awareness of political violence on social media is necessary to protect users. Not only do many graphic and politically charged posts go unchecked and pass through a platform's guidelines, but users can be stuck in echo chambers within the platform. Echo chambers give users a space to engage in conversations and consume information that typically reinforces their beliefs. These chambers personalize feeds tailored to the engagement and interactions of the user, which initiate feedback loops that bring in more content on the same topic. Based on the frequency and duration of engagement with each post, algorithms will suggest more content on the same topic through recommendation engines to amplify the exposure of politically violent content after a user initially interacts with a politically violent post. Such frequent and aggressive exposure can lead to emotional desensitization to political violence since uncensored, graphic posts become "normal" and no longer disturbing.

I. Literature Review

Although there is a lack of information regarding the relationship between political violence exposure on social media and emotional desensitization among American high school students, researchers in other countries have begun to investigate this phenomenon.

Baseline: American Students

There are a few studies that have been conducted using Americans or in the United States, such as those by the American Psychological Association, Brittany Bostic, A. Oshodi, and Erica Scharrer. Many of these studies are literature-based and focus on the findings from previously conducted studies using both American and international subjects. However, both A. Oshodi and Erica Scharrer conducted surveys and interviews with Americans to research the impact of exposure to media violence. Both found that increased violence exposure affected participants' responses to violence, typically increasing violent attitudes. Furthermore, they found that frequency and duration of exposure played a role in the impact of violence on participants. These studies served as a baseline of understanding regarding the relationship between Americans' exposure to political violence and desensitization.

In Other Countries

Researchers in countries other than America have begun researching this phenomenon. The results are consistent with expectations: a clear correlation between exposure to political violence on social media and emotional desensitization.

I. Colombia

Researcher Esteban Morales, a professor of Media and Digital Cultures at the University of Groningen, studied the relationship between Colombian young adults' social media usage and their reactions to everyday violence. His study revealed and emphasized how social media platforms impact the process of interpreting and justifying violence. However, his most critical finding was that violence-exposed young adults exhibited feelings of normalization to violence. Furthermore, Morales concluded that "violence becomes just *another* content that young adults encounter in their everyday engagements with social media" (Morales). Finally, he discovered that young adults' exposure to such violent social media content transforms their interactions with others and can lead to their decision to enact harm to others.

II. Egypt

In hopes of studying the role of social media in promoting peaceful interactions and preventing violence, researchers Maha El Tarabaishi and Samar Galal asked Egyptian university students to complete a questionnaire exploring their beliefs. Maha El Tarabaishi is a well-known researcher in Media and Communication at a university in Egypt. Samar Galal is a professor in Mass Communication and Journalism at October University for Modern Sciences and Arts. Although they found that social media can build peace and prevent violence, ultimately, the algorithm works against these advantages by pushing perspectives that align with each user. Therefore, a user who justifies violence will interact with other users who also justify violence.

III. Germany

Professor of Social Psychology at the University of Potsdam, Barbara Krahé, researched the connection between habitual social media violence exposure and desensitization to violence. Krahé found that individuals exposed to violent social media content experienced and demonstrated less physiological reactivity upon viewing such content. Although both women and men were desensitized to the content, women were significantly more desensitized than men.

IV. Iran

Dr. Mohammad Dalvi-Esfahani is a professor of Business Analytics at the University of Southampton Malaysia. He researched how social media affects empathy and personality traits using a cross-sectional correlation survey for Iranian adolescents. Dalvi-Esfahani found that individuals with low empathy levels use social media to compensate for their lack of belonging and real-world experiences, and therefore may develop social media addictions. Upon frequent and heavy use of social media, users can be exposed to violent content and continue to experience decreased levels of empathy.

V. Israel

Shira Dvir Gvirsman is a professor in Communication Studies at Tel Aviv University and has sought to determine the effects of political violence exposure on Israeli youth. Dvir Gvirsman gathered participants to take questionnaires and compared the data against political

violence, community violence, school violence, and family violence scales. She found that there was a direct correlation between social media violence exposure and reinforcing behavior difficulties. Although the participants demonstrated vulnerability to behaviors like aggression, they did not express strong emotional reactions to the content.

VI. Lebanon

Asma Tarabah is a professor in Pediatrics at the American University of Beirut and wanted to research children's desensitization in relation to exposure to violence. Tarabah found that the frequency and aggressiveness of exposure do impact desensitization. Children are especially at risk for long-lasting effects on their moral judgments and mental health.

VII. Nigeria

Professors in Politics at the University of Nigeria, Dorathy Isaac and Chieme Azubuike, researched the advantages and disadvantages of social media political participation. While they found that social media increases political participation, the study also revealed that social media political participation increases hate messaging and the spread of graphic, violent content. Such content was reinforced by algorithms that frequently suggested more images and videos to users. Isaac and Azubuike believe that the frequency of exposure would lead to desensitization to violence.

VIII. Pakistan

Noor Fatima is a political researcher who wanted to determine the role that social media plays in encouraging violent behaviors among Pakistani university students. Fatima found that social media can foster conformity, ultimately leading to violent conduct. She found that exposure to violence on social media caused aggression, frustration, and other behavioral concerns that normalize violent tendencies.

IX. Palestine

To determine the connection between political violence and mental health issues, Professor of Community Health at Al-Quds University Ziad Abdeen questioned Palestinian youth to measure their understanding and reaction to political violence. Abdeen found that exposure to political violence can lead to mental health decline and suicide attempts. However, youth who were frequently exposed to violence demonstrated no significant reaction. Abdeen concluded that political violence mainly affects youth behaviors, but not their emotions.

Discussion

Overall, each study found similar trends and findings across populations. In response to violent content exposure, a common finding was a lack of emotional reaction and decreased empathy (Morales, Krahé, Dalvi-Esfahani, Gvirsman, Fatima, Abdeen). With increased exposure

to violence, the respondents studied expressed no significant reactions and demonstrated low levels of empathy for the victims.

Furthermore, another major finding was that algorithms push content that aligns with a user's opinions and can recommend politically violent content on social media (El Tarabaishi, et al., Dalvi-Esfahani, Isaac, et al.). By increasing the amount of political violence consumed by a user, there is a greater risk of desensitization.

Finally, several studies found that exposure to violence on social media reinforces negative behavior, such as aggression, anger, or frustration (El Tarabaishi, et al., Dalvi-Esfahani, Gvirsman, Tarabah, Fatima).

Political Conflicts Bias

Although the studies above recorded consistent trends of desensitization from political violence posted on social media, it is important to recognize that the majority of the countries above have experienced ongoing politically violent conflicts. Such events could have influenced the responses of survey and interview participants in the above studies, effectively skewing the data. For instance, Colombian citizens experience cartel violence, and Nigerian citizens experience religious persecution from the terrorist group Boko Haram. Similarly, the Israeli-Palestinian conflict, an ongoing war for over a decade, impacts Israel, Palestine, and nearby countries, such as Egypt, through the Rafah Border Crossing. The Russia-Ukraine War, ongoing since early 2022, may affect data found in studies with populations from countries like Germany due to its proximity to Russia and Ukraine. These groups, events, and wars may invoke political desensitization due to personal and direct exposure to political violence.

Similar to the studies above, this study uses a population that has been directly and personally impacted by such ongoing political conflicts. Therefore, the responses found in this study may be influenced by causes beyond social media. It is important to note that social media may not be the only cause of the reactions and responses reported by survey respondents when analyzing the findings.

Gap

While this phenomenon has been heavily studied and researched in foreign countries (Abdeen et al.; Azubuike & Isaac; Dalvi-Esfahani; El Tarabaishi & Galal; Fatima et al.; Gvirsman; Krahé et al.; Morales; Tarabah et al.), there is a lack of information regarding how social media political violence exposure affects Americans.

Currently available studies mostly focus on children in elementary or middle schools, college students, or young adults, ultimately skipping high school students (Abdeen et al.; Azubuike & Isaac, Dvir Gvirsman et al.; El Tarabishi & Galal; Fatima et al.; Krahé et al.; Morales; Tarabah et al.).

Although there is limited information and studies available regarding American high school students. They are increasingly exposed to social media at a young age, making them vulnerable to emotional desensitization. Thus, the researcher was inspired and interested in

focusing on American high school students in this study of the relationship between exposure to political violence on social media and desensitization.

Furthermore, current American political violence events, such as the death of Charlie Kirk, inspired the researcher to explore this phenomenon. After seeing the aftermath of such a politically violent event, the researcher recognized the importance of understanding how that type of content impacts American high school students.

II. Method

To analyze the relationship between politically violent content on social media and desensitization among American high school students, the researcher utilized a mixed-methods research approach. Based on the methods used in studies such as those by Arias et al. and Martínez et al., the researcher distributed a survey to local high school students and conducted interviews with child psychologists. The researcher determined that this approach would be most effective because of the perspectives from experts and high schoolers. By collecting and analyzing both qualitative and quantitative data, the researcher believed the data would be stronger and reveal clearer trends to help fill the gap in research on American high school students.

Population

This study involved both a survey of local high school students and interviews with child psychologists. The survey was distributed to high schools near Calabasas, California. Each grade at each high school had the opportunity to participate in the survey. Additionally, the researcher reached out to over two hundred child psychologists in Los Angeles County, but only received a handful of responses. To still include the perspective of child psychologists, the researcher conducted interviews with child psychologists at Calabasas High School. By using both personal perspectives from high school students and expert opinions from local child psychologists, the researcher hoped to identify and understand patterns among the students.

Instruments

The survey was developed using questions and ideas from surveys used in other studies. The vignettes, or hypothetical situations, were used to understand students' thinking processes when exposed to political violence. The researcher chose to distribute the survey digitally to increase accessibility across participating high schools.

Survey Questions

Survey question	Source the question is based on
What grade are you in?	Influence of Social Media in Shaping the Adolescents' Political Awareness, Participation, and Activism

What is your gender?	Influence of Social Media in Shaping the Adolescents' Political Awareness, Participation, and Activism
What high school do you attend?	Influence of Social Media in Shaping the Adolescents' Political Awareness, Participation, and Activism
How would you describe your political interest/participation? (Likert Scale)	Social Media: A Breeding Ground for Energy or Apathy?
What social media platforms do you use frequently? Check all that apply.	Exposure to Violence and Children's Desensitization Attitudes in Lebanon
Do you see any of the following in your feed? Check all that apply.	Self-defined
On average, how many hours do you spend on social media per day?	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
When did you get your first form of social media?	Self-defined
In the past 6 months, how often have you seen posts, videos, or comments on social media that show real political violence (ex: fights, weapons, attacks)? (Likert Scale)	Students' Exposure to Social Media and Their Radical Involvement on the Societal Issues in the Philippines
In the past 6 months, how often have you seen posts that *encourage* or *praise* violence for political reasons? (Likert Scale)	Students' Exposure to Social Media and Their Radical Involvement on the Societal Issues in the Philippines
How often do you look for political content on social media? (Likert Scale)	Evaluation of Social Media and Students' Participation in Politics: A Study of Students in the University of Nigeria, Nsukka
How often do you share or repost political posts that show or praise violence? (Likert Scale)	Evaluation of Social Media and Students' Participation in Politics: A Study of Students in the University of Nigeria, Nsukka
Democrats are more likely to commit acts of political violence.	Social Media: A Breeding Ground for Energy or Apathy?
Republicans are more likely to commit acts of political violence.	Social Media: A Breeding Ground for Energy or Apathy?
How shocked did you feel? (Likert Scale)	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior

How sad did you feel? (Likert Scale)	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
How scared did you feel? (Likert Scale)	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
How angry did you feel? (Likert Scale)	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
How uninterested did you feel? (Likert Scale)	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
Seeing violent political posts makes political violence seem more normal.	Enhancing online safety: The impact of social media violent content and violence among teens in Illinois
When I see a person hurt in a political video, I feel sorry for them	Enhancing online safety: The impact of social media violent content and violence among teens in Illinois
Vignette	**Vignette**
A young man travels across America to debate with locals regarding political opinion. As he reached California, the young man was shot and killed due to another disagreeing with his opinion. Do you believe that the young man's death was justified? Do you believe that the man who shot him was justified?	Self-defined
An American President was arguing in favor of taking away civil rights, such as our First Amendment, which grants free speech. During a meeting, the President was stabbed and killed. In your opinion, was the President's death justified? Was the person who stabbed him justified?	Self-defined

The interviews were conducted using a list of ten prepared questions, based on questions used in interviews in other studies.

Interview Questions

Interview Question	The source that the question is based on
What behavioral or emotional indicators suggest that a teenager is becoming desensitized to violence?	Self-defined

Why are teenagers and adolescents vulnerable to desensitization? What aspects of the developing brain cause this, if any? Does social media take advantage of these aspects more than other platforms (such as the news)?	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
To what extent does peer social media interaction (comment sections, chat groups, friends reposting content on stories) impact desensitization? Do you believe a mob mentality or similar phenomenon is occurring?	Influence of Social Media in Shaping the Adolescents' Political Awareness, Participation, and Activism
Do you think that repeated politically violent content exposure can alter a student's moral or political reasoning? (Ex: believing violence is necessary or deserved)	Exposure to Violence and Children's Desensitization Attitudes in Lebanon
In your experience, how do adolescents typically react emotionally to politically violent content? How do adolescents overly exposed to such content react in comparison?	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
Is there a difference in emotional reaction from political violence exposure to general violence exposure?	Desensitization to Media Violence: Links With Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior
Are there lasting impacts of political violence exposure as an adolescent? Into adulthood?	Self-defined
Is there a greater impact of desensitization if the viewer is younger or older? Who is more susceptible: an elementary school, middle school, or high school student?	Self-defined
How do you think memes, humor, or edited videos of political violence impact the seriousness with which teens perceive such acts?	Influence of Social Media in Shaping the Adolescents' Political Awareness, Participation, and Activism

Sample Selection

The researcher asked to distribute the survey to the following high schools: Calabasas High School, Agoura High School, Westlake High School, Chaminade College Preparatory, Sierra Canyon High School, and Viewpoint School. However, most did not distribute the survey due to the researcher being an out-of-district student. Calabasas High School, Agoura High School, and Sierra Canyon High School were the only schools that distributed the survey. Six respondents came from Agoura High School and Sierra Canyon High School. The remaining 126 responses came from Calabasas High School. About 30 students from each grade answered the survey, except for the sophomores, who accounted for nearly 60 responses. The researcher chose a random probability sample selection because the survey was specific to high school students,

but no other trait. Moreover, a random probability sample would give each of the participants the same chance to participate, even if they went to different high schools.

The researcher contacted over two hundred child psychologists at random to gather participants for interviews. The researcher utilized a convenience sampling method to gather potential interview participants due to a lack of personal contacts. Additionally, the expert knowledge needed from these psychologists made the use of probability sampling difficult. While the researcher reached out to many psychologists, only two responded and were unable to set up a date to meet. In an effort to combat this, the researcher interviewed child psychologists who work for Calabasas High School.

Ethics

To conduct this study ethically, the researcher protected participants' anonymity and included informed consent forms. The researcher chose not to collect the names or emails of student survey participants in hopes of more honest responses and to protect their anonymity. Students were also asked to sign an informed consent form using their initials before beginning the survey. Furthermore, the interview participants' names were never recorded, and they were referred to using the chronological order of their interviews. These participants were also asked to sign an informed consent form before the interview began.

Implementation

To effectively distribute the survey to local high schools, the researcher contacted the principals at each school and requested that the survey be sent out to students. However, most did not distribute the survey due to the researcher being an out-of-district student. At Calabasas High School, the researcher reached out to teachers at random and asked them to distribute the survey through Google Classroom.

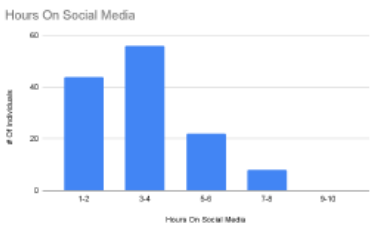
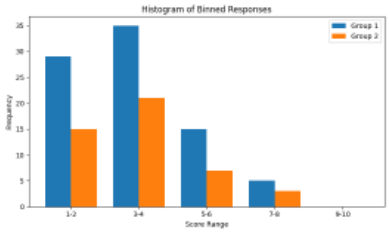
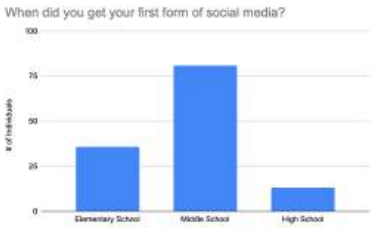
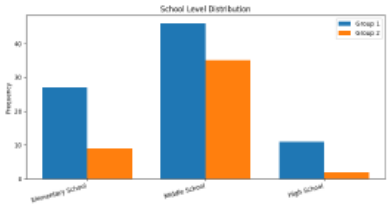
To schedule interviews, the researcher reached out to over two hundred child psychologists and received responses from only a few. Instead of using Los Angeles Child Psychologists, the researcher utilized Calabasas High School wellness counselors, who were more readily available and accessible for the researcher to meet with. These counselors worked with the majority of the students who completed the survey. The interviews were used to help analyze the data and trends found in the surveys against the child psychologists' expert knowledge.

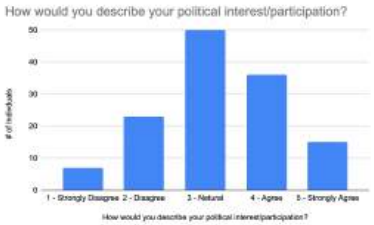
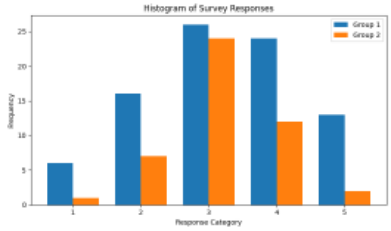
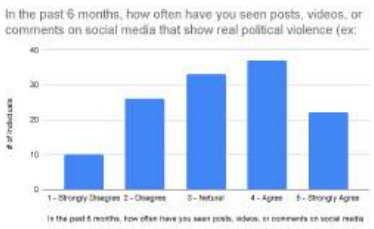
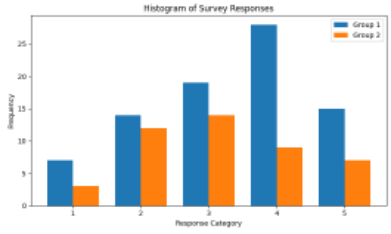
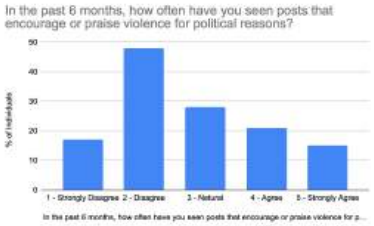
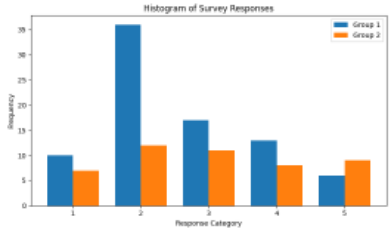
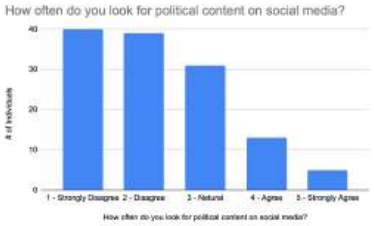
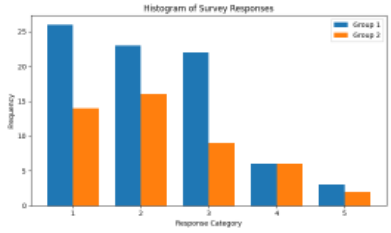
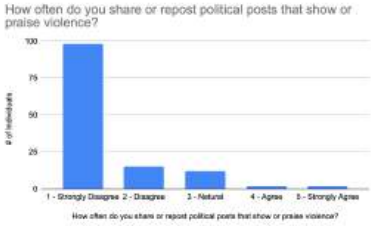
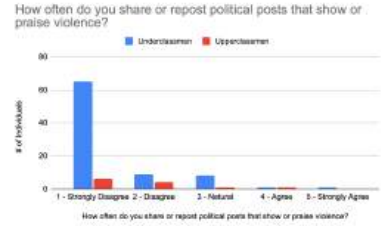
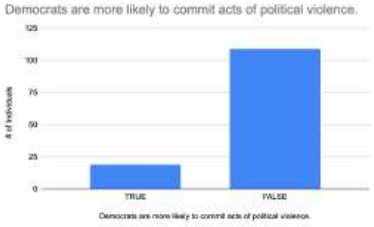
III. Findings & Analysis

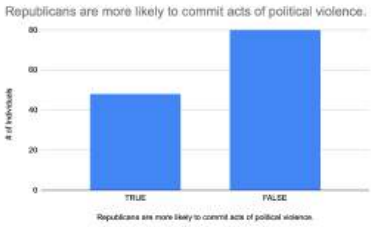
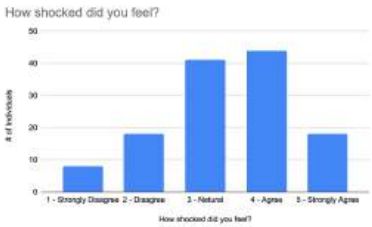
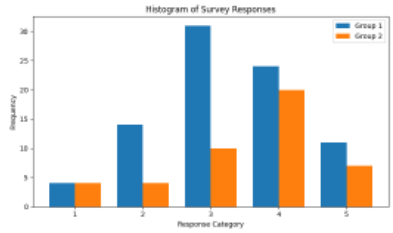
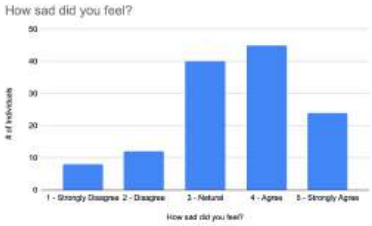
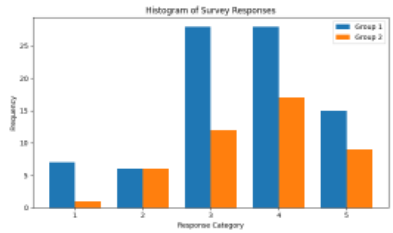
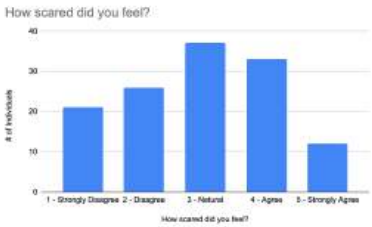
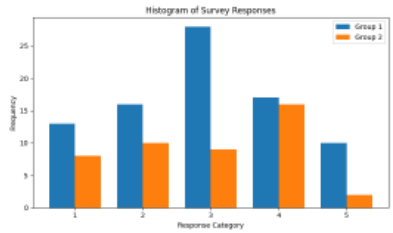
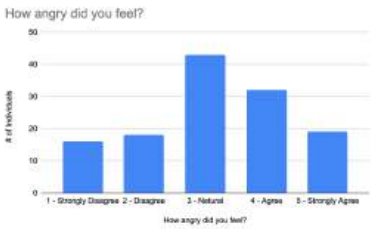
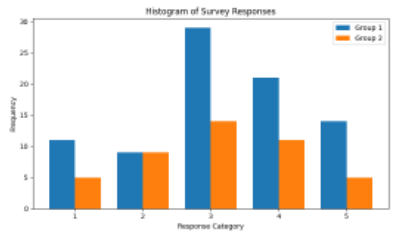
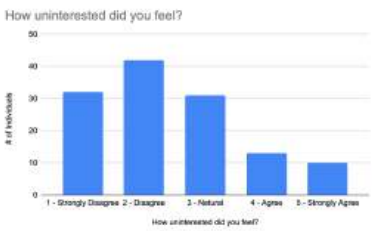
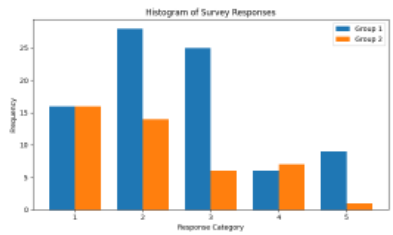
The data was collected and analyzed with the following three themes in mind, based on their presence found in literature and past studies (American Psychological Association, Arias et al., Conathan, Dalvi-Esfanhai, Hughes, Ibardeloza, Isaac et al., King, Li, Martinez et al., Montalbo, Scharrer):

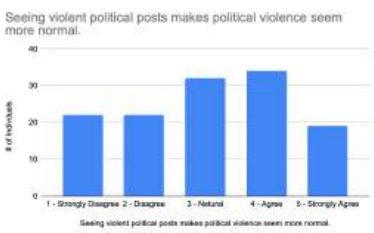
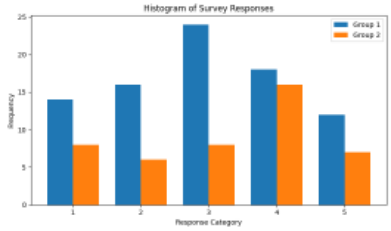
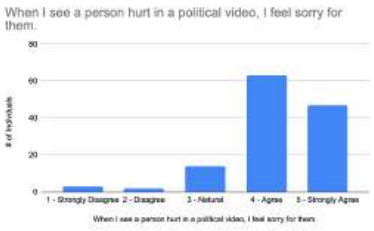
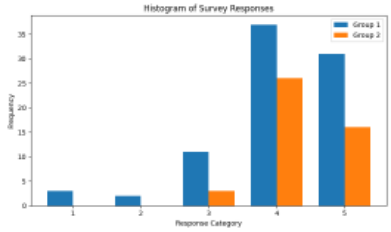
Theme	Definition
Lack of empathy	The inability to recognize, understand, and experience another person's emotions (Scharrer, Erica. "Media exposure and sensitivity to violence in news reports: Evidence of desensitization?" *Journalism & Mass Communication Quarterly* 85.2 (2008): 291-310.).
Emotional reaction	The natural reaction to a stimulus (Li, Jianing, Devin Conathan, and Ceri Hughes. "Rethinking emotional desensitization to violence: Methodological and theoretical insights from social media data." *Proceedings of the 8th International Conference on Social Media & Society*. 2017).
Political interest/involvement	Engagement and participation in societal and governmental issues (Arias, J., and L. Montalbo. "Influence of social media in shaping adolescents' political awareness, participation, and activism." *Asia Pacific Journal of Management and Sustainable Development* 12.2 (2024): 166-175).

The themes stated above were then used to develop questions on the student survey. The following data was collected through the student survey. The researcher analyzed the responses by grade level to examine the difference in trends among the ages surveyed. 9th and 10th grade students are considered "underclassmen", while 11th and 12th grade students are considered "upperclassmen".

Question #	All Grades (A)	Underclassmen vs. Upperclassmen (B)
1.		
2.		

3.	 How would you describe your political interest/participation? # of Individuals 1 - Strongly Disagree 2 - Disagree 3 - Natural 4 - Agree 5 - Strongly Agree How would you describe your political interest/participation?	 Histogram of Survey Responses Group 1 Group 2 Frequency Response Category
4.	 In the past 6 months, how often have you seen posts, videos, or comments on social media that show real political violence (ex: # of Individuals 1 - Strongly Disagree 2 - Disagree 3 - Natural 4 - Agree 5 - Strongly Agree In the past 6 months, how often have you seen posts, videos, or comments on social media	 Histogram of Survey Responses Group 1 Group 2 Frequency Response Category
5.	 In the past 6 months, how often have you seen posts that encourage or praise violence for political reasons? % of Individuals 1 - Strongly Disagree 2 - Disagree 3 - Natural 4 - Agree 5 - Strongly Agree In the past 6 months, how often have you seen posts that encourage or praise violence for p...	 Histogram of Survey Responses Group 1 Group 2 Frequency Response Category
6.	 How often do you look for political content on social media? # of Individuals 1 - Strongly Disagree 2 - Disagree 3 - Natural 4 - Agree 5 - Strongly Agree How often do you look for political content on social media?	 Histogram of Survey Responses Group 1 Group 2 Frequency Response Category
7.	 How often do you share or repost political posts that show or praise violence? # of Individuals 1 - Strongly Disagree 2 - Disagree 3 - Natural 4 - Agree 5 - Strongly Agree How often do you share or repost political posts that show or praise violence?	 How often do you share or repost political posts that show or praise violence? Underclassmen Upperclassmen # of Individuals 1 - Strongly Disagree 2 - Disagree 3 - Natural 4 - Agree 5 - Strongly Agree How often do you share or repost political posts that show or praise violence?
8.	 Democrats are more likely to commit acts of political violence. # of Individuals TRUE FALSE Democrats are more likely to commit acts of political violence.	NA

9.		NA
10.		
11.		
12.		
13.		
14.		

15.		
16.		

The researcher also conducted a few interviews with child psychologists to help contextualize the trends demonstrated in the survey. The researcher interviewed a child psychologist working at Calabasas High School. The child psychologists interviewed have met with the majority of survey participants. In the following chart, the interview is summarized by the points made by the participant after being asked a question and the theme that correlates with their response. Some responses relate to more than one theme and yield multiple important results.

The interview participants had several overlapping ideas and responses to questions. Each participant agreed that a lack of empathy or any response to the political violence is the first indicator of desensitization. They all related the lack of a response to the brain and critical thinking skills in high schoolers being underdeveloped. Furthermore, each participant responded that political affiliations and opinions would be reflective of those of parents or friends. They attributed this trend to a lack of self-esteem and confidence in high school students increasing their desire to belong and not ‘stand out’, which causes many students to carry similar political perspectives. Social media was noted to validate opinions due to the ability to be an echo chamber and only be presented with opinions that reinforce the ones currently held. Finally, each participant mentioned that politically violent content posted on social media reduces accountability and creates a tolerance for violence. Viewers typically have little to no direct connection to such acts, causing a lack of response to the stimulus.

Participant	Summary of Interview	Theme
Participant 1	1. No response to or avoiding violence; fatigue of viewing violence 2. Brains are underdeveloped; easier to avoid violence than to try to understand; parents stop protecting	1. Emotional reaction; political interest 2. Emotional reaction; Political interest

	children from violence as they get older (increased exposure); more accessible; includes opinions; students seek information on these platforms more 3. Self-esteem issues and still forming opinions; beliefs reflect those of students' parents; picking a side and sticking with it (viewpoint) 4. Becomes a norm; negative connotation with the news/news reports 5. Teens follow behavior; students may feel differently depending on their political ideas 6. Violence is heightened when political or racial; reactions are greater for political violence rather than typical violence 7. Exposure to violence at a young age causes the formation of opinions earlier than necessary; forced into making political choices earlier in life; hinders the ability to see both sides 8. High school students are less protected from violence; younger youth have questions 9. Desensitization increases when violence is made to be humorous; not funny when the violence becomes personal; it separates one from the issues	3. Lack of empathy; Political interest 4. Lack of empathy; Political interest 5. Political interest; Emotional reaction 6. Emotional reaction 7. Political interest; Emotional reaction 8. Emotional reaction 9. Lack of empathy
Participant 2	1. Witnesses a lack of empathy; the victim is seen as less than human; violent behavior 2. The brain is underdeveloped; social media takes advantage of the neuroplasticity of the teenage brain; exposure therapy helps combat phobias with repeated exposure 3. Identity within politics; aligning with opinions that are not fully ours; desire for belonging or connection 4. Can be hard to see both sides of an argument; students realize they can have an impact 5. Internalize it; mental health concerns and confusion handling emotions; increased defensiveness or looking out for danger 6. Not much of a difference; it affects the brain in a similar way 7. Possibility of radicalization; opinions are changed; impact on future generations 8. Impact is larger when students are younger, different at different ages 9. Dehumanization; takes away accountability	1. Lack of empathy 2. Emotional reaction 3. Political interest 4. Political interest 5. Emotional reaction 6. Emotional reaction 7. Emotional reaction 8. Emotional reaction; Political interest 9. Lack of empathy
Participant 3	1. Lack of empathy in their response to a situation 2. Exposure to aggressive violence through the internet; News has a responsibility to fact-check and be ethical, which social media does not have; social media	1. Lack of empathy 2. Emotional reaction; Political interest 3. Lack of empathy

	validates opinions 3. Dehumanization; increased polarization 4. Easy to be in an echo chamber; see political violence as a call to action or a need to retaliate 5. Critical thinking is underdeveloped in youth; easier manipulation of younger students; potential to experience more extreme reactions as a younger adolescent 6. Political violence is tied to political ideology; politics can become an identity 7. Exposure to anything that causes the formation of an opinion will be carried through the rest of life; political violence is a trauma 8. High school students would be more susceptible due to increased independence; figuring things out for themselves 9. Minimizes the violent act; creates tolerance for violence	4. Political interest; Emotional reaction 5. Emotional reaction 6. Political interest; Lack of empathy 7. Emotional reaction 8. Emotional reaction 9. Lack of empathy

As seen in the graphs, underclassmen spend more time on social media and get access to social media later than upperclassmen, on average (Figure 1B and Figure 2B). Based on the information gathered during interviews with the child psychologists, students begin developing their own opinions as they age. Therefore, getting social media at a later age and spending more time on such platforms may cause underclassmen to be more opinionated on political topics and use social media as an outlet to engage in political matters. Aligning with this idea, underclassmen demonstrated greater political interest and participation than upperclassmen (Figure 3B and Figure 6B). Underclassmen also seek political content, such as news, on social media, and reported more exposure to politically violent content within the last six months than upperclassmen (Figure 4B, Figure 5B, Figure 6B).

While the political engagement among underclassmen is greater than that of upperclassmen, their reactions are also greater. Many report feelings of empathy, fear, anger, sadness, and shock when viewing politically violent content (Figure 10B, Figure 11B, Figure 12B, Figure 13B, Figure 14B, and Figure 16B). Participant 3 noted that younger students may experience a more intense or extreme reaction to the content due to their underdeveloped brains and critical thinking skills. Participant 1 also mentioned that political violence causes greater reactions, as opposed to ‘regular’ violence.

However, underclassmen agreed that viewing politically violent content normalizes it, despite their intense reactions (Figure 15B). In their interview, Participant 2 related political violence exposure and desensitization to exposure therapy. As defined by researchers Strickland, Hodges, North, and Weghorst, exposure therapy subjects a patient with a phobia to real-life, anxiety-inducing stimuli, in hopes of normalizing their fear (Strickland et al). Therefore, since the underclassmen reported intense emotional reactions while experiencing the normalization of political violence, they may have limited past exposure to political violence. While they agreed

that it normalizes it, they are not accepting of political violence based on their reactions. Repeated exposure to content that induces intense emotional reactions can lead to desensitization, as it becomes normal to viewers.

IV. Discussion

The survey and interviews were tailored to answer three themes: lack of empathy, emotional reaction, and political interest and involvement. These themes were the most present and significant across previous research (American Psychological Association, Arias et al., Conathan, Dalvi-Esfanhai, Hughes, Ibardeloza, Isaac et al., King, Li, Martinez et al., Montalbo, Scharrer).

Lack of Empathy

Defined by Erica Scharrer as the inability to recognize, understand, and experience another person's emotions, interview participants 2 and 3 both agreed that a lack of empathy is the initial sign they observe in desensitized individuals. However, the majority of student-survey respondents recorded feelings of empathy for victims of political violence (Figure 15A). Such trends are a direct contrast to studies such as the American Psychological Association and Mohammad Dalvi-Esfanhai, both of which found that high media exposure to political violence and social media addiction decreased empathy. Instead, this study's results were consistent with a study conducted by Erica Scharrer, who found that people with higher levels of empathy showed stronger emotional responses to political violence.

Emotional Reaction

Defined as the natural reaction to a stimulus by authors Jianing Li, Devin Conathan, and Ceri Hughes, students expressed significant emotional reactions to political violence media exposure, which is consistent with the data found by Erica Scharrer. In past studies, such as the ones conducted by Jianing Li et al. and Vania Martinez et al., emotional response intensity decreased with significant violence exposure.

Once again, these trends were not consistent with the data found in the student survey in this study (Figure 10A, Figure 11A, Figure 12A, Figure 13A, Figure 14A). However, the type of violence studied differs between these studies, as those tested general violence, and this one researched political violence specifically. As explained by interview participant 1, political or racial violence has a greater potential for intense reactions, rather than general violence, which may explain the difference in emotional reactions.

Political Interest and Involvement

As defined by researchers Arias and Montalbo, political interest and involvement is the engagement and participation in societal and governmental issues. The majority of the student survey respondents reported slight interest and involvement in politics (Figure 3A, Figure 4A, Figure 5A, Figure 6A).

Past studies, like Dorathy Isaac et al., J. Arias et al., and Liana King, have attributed increased political interest and involvement to social media. They found that social media was reported as a tool to communicate with others, express their opinions, and call to action. However, Kenneth B. Ibardeloza found that social media was used for awareness, as opposed to activism.

Although some students did report using social media to increase their political engagement and to stay updated on the news, most reported that they do not use social media to further their political involvement or interest.

V. Conclusion

This study analyzed American high school students' desensitization to political violence on social media using three themes: lack of empathy, emotional reaction, and political interest/involvement. Although most studies found that increased political violence exposure led to decreased emotional reaction, this study discovered that increased political interest and involvement caused greater emotional reactions and empathy. Many students reported that they felt political violence exposure does normalize it, which contrasts with their reactions to it. This study helps fill the gap by providing insight into the relationship between political violence exposure on social media and emotional desensitization in American youth.

Implications

Youth desensitization to political violence due to social media exposure is a developing issue. These findings suggest the need for digital literacy educational programs to encourage healthy boundary-setting online and protection from excessive media consumption that leads to desensitization.

Additionally, these programs can help teachers meet new state digital literacy educational standards. In California, Assembly Bill 2876 was passed in 2024, which required the Instructional Quality Commission to consider including media literacy content in its criteria. Since then, existing curriculum standards have incorporated select programs to educate students on digital literacy skills (California Department of Education). Following California, other states should adopt similar standards and programs to encourage students to use social media safely.

This study also suggests the need for emotional counseling, so students have the ability to properly respond to political violence. Such practices can support students to ensure that they understand the content they are viewing while maintaining emotionally healthy responses.

Ultimately, this study demonstrates the need for youth protection on social media.

Limitations & Future Studies

While this research offers a better understanding and insight into youth desensitization, there are a few limitations. The sample size of the survey was limited mainly to Calabasas High School, with a few responses from other local high schools, ultimately limiting perspectives. Furthermore, the survey was taken by 50% more sophomores than any other grade, which could

have impacted the data and trends depicted by the responses. To combat this, the researcher analyzed data by percentage instead of by number of participants. Finally, the sample size of the interviews was relatively small and geographically limited to Calabasas High School, further inhibiting diversity in opinions or viewpoints.

Furthermore, the survey contained self-reported answers, which may be subject to social desirability bias and conformity. Students do not want to stand out and may feel uncomfortable answering the questions honestly. In the interview with Participant 1, they mentioned that students experience self-esteem issues at this age and may side with opinions that reflect those of their friends or parents.

Another significant limitation to acknowledge is the direct and personal ties to ongoing political conflicts, specifically the Israeli-Palestinian conflict. Such a connection to an ongoing conflict would increase exposure to political violence, which may affect the student survey responses.

For future studies and research, a broader sample size and location of participants would help gain more insight into trends and ensure that more perspectives are accounted for. Additionally, a study comparing elementary, middle, and high school students would be beneficial for future research to determine which age groups are most at risk of becoming desensitized.

Ultimately, expanding research across geographical areas and ages will help provide a deeper understanding of how this issue may develop and how to combat it.

Works Cited

Abdeen, Ziad, et al. "The Association between Political Violence and the Connection between Bullying and Suicidality among Palestinian Youth." Suicide and Life-Threatening Behavior, vol. 48, no. 1, 2018, pp. 95-104.

American Psychological Association. "Violence in the Media Psychologists' Study Potential Harmful Effects." American Psychological Association, 2013.

Arias, J., and L. Montalbo. "Influence of Social Media in Shaping Adolescents' Political Awareness, Participation, and Activism." Asia Pacific Journal of Management and Sustainable Development, vol. 12, no. 2, 2024, pp. 166-175.

Bostic, Brittany. "Does Social Media Perpetuate Youth Violence? - Michigan Youth Violence Prevention Center." Michigan Youth Violence Prevention Center, 20 Feb. 2014, yvpc.sph.umich.edu/social-media-perpetuate-youth-violence/?highlight=does%20social%20media.

Dalvi-Esfahani, Mohammad, et al. "Social Media Addiction and Empathy: Moderating Impact of Personality Traits among High School Students." Telematics and Informatics, vol. 57, 2021, p. 101516.

Dvir Gvirsman, Shira, et al. "The Effects of Mediated Exposure to Ethnic-Political Violence on Middle East Youth's Subsequent Post-Traumatic Stress Symptoms and Aggressive Behavior." Communication Research, vol. 41, no. 7, 2014, pp. 961-990.

ElTarabishi, Maha, and Samar Galal. "The Role of Social Media in Promoting Peace Building and Prevention of Societal Violence among University Youth (Analytical Descriptive Study)." Journal of Media and Interdisciplinary Studies, vol. 2, no. 4, 2023, pp. 1-32.

Fatima, Noor, et al. "The Role of Social Media in Proliferating Violent Behavior Among College Students in District Layyah, Pakistan." Review of Applied Management and Social Sciences, vol. 5, no. 4, 2022, pp. 525-537.

Ibardeloza, Kenneth B., et al. "Students' Exposure to Social Media and." International Review of Social Sciences Research, vol. 2, no. 1, 2022, p. 47.

Isaac, Dorathy, and Chieme Azubuike. "Evaluation of Social Media and Students Participation in Politics: A Study of Students in the University of Nigeria, Nsukka." European Scientific Journal, vol. 15, no. 20, 2019, pp. 1-17.

King, Liana. "Social Media: A Breeding Ground for Energy or Apathy?" Perceptions, vol. 5, no. 1, 2019.

Krahé, Barbara, et al. "Desensitization to Media Violence: Links with Habitual Media Violence Exposure, Aggressive Cognitions, and Aggressive Behavior." Journal of Personality and Social Psychology, vol. 100, no. 4, 2011, p. 630.

Li, Jianing, Devin Conathan, and Ceri Hughes. "Rethinking Emotional Desensitization to Violence: Methodological and Theoretical Insights from Social Media Data." Proceedings of the 8th International Conference on Social Media & Society, 2017.

"Media Literacy Resources." Media Literacy Resources - Curriculum and Instruction Resources (CA Dept of Education),

www.cde.ca.gov/ci/cr/ml/#:~:text=Understanding%20and%20teaching%20media%20literacy,integrate%20information%20into%20their%20lives. Accessed 29 Apr. 2026.

Martínez, Vania, Álvaro Jiménez-Molina, and Mónica M. Gerber. “Social Contagion, Violence, and Suicide among Adolescents.” Current Opinion in Psychiatry, vol. 36, no. 3, 2023, pp. 237-242.

Morales, Esteban. “Social Media and the Mediation of Everyday Violence: A Study of Colombian Young Adults’ Experiences.” New Media & Society, vol. 27, no. 7, 2025, pp. 4308-4325.

Mrug, Sylvie, et al. “Emotional Desensitization to Violence Contributes to Adolescents' Violent Behavior.” Journal of Abnormal Child Psychology, vol. 44, no. 1, 2016, pp. 75-86.

Oshodi, A. N. “Enhancing Online Safety: The Impact of Social Media Violent Content and Violence among Teens in Illinois.” World Journal of Advanced Research and Reviews, vol. 23, no. 3, 2024, pp. 826-833.

Patton, Desmond Upton, et al. “Social Media as a Vector for Youth Violence: A Review of the Literature.” Computers in Human Behavior, vol. 35, 2014, pp. 548-553.

Scharrer, Erica. “Media Exposure and Sensitivity to Violence in News Reports: Evidence of Desensitization?” Journalism & Mass Communication Quarterly, vol. 85, no. 2, 2008, pp. 291-310.

Tarabah, Asma, et al. “Exposure to Violence and Children’s Desensitization Attitudes in Lebanon.” Journal of Interpersonal Violence, vol. 31, no. 18, 2016, pp. 3017-3038.

Two Generations, One Amendment: Elizabeth Cady Stanton, Alice Paul, and the Era-Dependent Tactics of the American Woman Suffrage Movement

By Sarina Singhal

Abstract

Elizabeth Cady Stanton and Alice Paul are remembered as two of the most influential figures in the American campaign for women suffrage during the 19th and 20th century, yet they belonged to different generations, held alternate ideologies, and instituted vastly different leadership strategies. This paper compares the two leadership styles, effectiveness and the role each person contributed to the suffrage movement's long-term success. Drawing on their own writings, the records of the organizations they led, and the work of later historians, it argues that Stanton's and Paul's sharply different methods—Stanton's foundational work of persuasion, ideology, and broad reform in the nineteenth century, and Paul's militant, single-issue confrontation in the 1910's—were each influenced by and suited to the distinct needs of their respective eras. Measured by lasting contribution rather than immediate legislative victory, neither leader can be ranked simply above the other; instead, their contrasting approaches show that successful social movements require different styles of leadership at different stages of development. Further, this paper proposes that the diverse approaches from these contrasting leaders were both complementary and essential to achieve reform.

Introduction

In the saga of the American women's rights movement, Elizabeth Cady Stanton and Alice Paul stand as icons in the struggle to obtain the right to vote. Although they never met, their lives intersected in a critical way: the passing of a generational torch. When Stanton died in 1902, Paul was just a teenager, yet Stanton's work had a profound effect on Paul's upbringing.[1] Paul's mother, Tacie Paul, was a staunch suffragist, and Paul grew up surrounded by the discussions and decisions of the National American Woman Suffrage Association (NAWSA).[2] Within a few years of Stanton's passing, Paul was actively involved in the movement herself. By the time the Nineteenth Amendment was ratified under Paul's leadership, over seventy years had passed since Stanton's 1848 Seneca Falls Convention.[3] These two leaders remain the defining pillars of the movement—Stanton at the beginning to construct the framework for the cause, and Paul at its conclusion to orchestrate its victory.

This generation gap is why comparing them as leaders becomes essential in analyzing the suffragist movement. They did not just live in different eras; they held different ideologies, pursued different goals, and instituted radically different strategies. Stanton lived in the mid-1800s and was an orator, writer, and theorist. She produced publications and delivered speeches with the goal of achieving broad social reform in all areas of gender inequality, including marriage laws, property rights, and religious representation, in addition to the right to vote. Paul lived in the early 1900s and was a strategist and organizer. She used a radical, militant

approach to focus on the singular goal of achieving a constitutional amendment granting women the right to vote.

Because there is such a profound difference in their goals and tactics as leaders in the suffrage movement, the question of "who was more effective?" is difficult to answer. If the definition of "effective" means who achieved legislative success, then the superficial answer would be Alice Paul, who was a pivotal leader when the Nineteenth Amendment was ratified. If instead the definition of "effective" is who created lasting change in the movement, then both leaders emerge as not only equally effective, but equally essential. By leading the movement in their respective eras with different ideologies and contrasting strategies, they both provided not just what the movement needed, but delivered it at the precise moment in history that the movement needed to achieve lasting social change.

Two Generations, One Movement

The first step in understanding why Elizabeth Cady Stanton and Alice Paul pursued different goals with widely disparate tactics requires an understanding of the era in which each one lived and the environment in which they were raised. Stanton was born in 1815 to a wealthy judicial family. Exposed to her father's law practice, she became aware at a young age of the disparate laws that existed between men and women and, more importantly, that the path to change relied on changing those laws.[4] Despite her privileged upbringing in a liberal family which allowed her to be educated beyond the usual level for women of that era, she still felt the limits of her gender. She was unable to attend most universities, was given narrow access to employment, and continued to have highly restricted legal rights; consequently, she remained keenly aware of the gender inequality that existed across all aspects of society.

It was upon her marriage, where she famously redacted the word "obey" from her vows, that she first became closely involved in social reform. Her husband was an abolitionist, and the anti-slavery movement was her first active involvement in a movement for social change.[5] Her honeymoon occurred in London during the World Anti-Slavery Convention where her husband was a speaker. There, she and the other women, including noted reformers such as Lucretia Mott, were denied seats on arrival to the convention. The uproar brought the convention to a halt, and after a day of heated debate, the women were reluctantly allowed to attend under the condition that they remain seated quietly behind a curtain. The humiliation from that experience, combined with meeting Mott behind that curtain, planted the idea of a convention dedicated to the rights of women. Years later, living as a young mother in Seneca Falls and feeling like a "caged lioness", she again had the opportunity to rekindle that conversation with Lucretia Mott and that fledgling idea was converted to organized action.[6]

When Stanton began her work in the mid-19th century, women were, as Stanton described it, "civilly dead". Under the legal doctrine of coverture, a married woman lost all legal identity to her husband: she could not own property, sign contracts, keep her own wages, or even keep her children if they separated. Her husband had the right to own her, even to the point of legally being able to beat her. Unmarried women, although considered independent individuals, faced

many obstacles including social stigma, limited educational opportunities, and few economic roles. The immense social and economic pressure led to marriage rates exceeding 90%, usually by the time women were in their early 20s.[7] Those that remained single were highly unlikely to have attended university and were limited to low-paying jobs such as textile workers, seamstresses, and laundresses, where they earned a fraction of the rate their male counterparts were paid.

Stanton's reform agenda at the Seneca Falls Convention in 1848 was as broad as the rights women were denied. The very wording of that document, which deliberately held an echo of the Declaration of Independence—"We hold these truths to be self-evident: that all men and women are created equal"— both highlighted the hypocrisy of the position of the American woman and declared her intention of obtaining complete equal rights for women.[8] The Declaration of Sentiments listed grievances ranging from unfair property laws, restrictive marriage and divorce laws, decreased access to education and employment, changes in religious representation, and ultimately, to the right to vote. Surprisingly, the call for women's right to vote was the most controversial of the demands. Many of the convention's delegates, including her own husband, were concerned that it was too radical and would make their attempts at reform look ridiculous. Only after Frederick Douglass gave an impassioned speech in support of suffrage did the delegation finally vote to include it in the Declaration of Sentiments.[9]

The focus of the movement slowly shifted over the next twenty years. The primary force for change occurred when the 14th and 15th amendments were proposed to extend citizenship and voting rights to Black men. The active exclusion of women split the movement. Stanton refused to support the amendments if women were excluded, while other activists supported the federal legislature, choosing to pursue voting rights on a state level instead.[10] This fractured the suffrage movement into two factions—Stanton's National Woman Suffrage Association and the American Woman Suffrage Association.[11] It took two decades for the movement to unite to become the National American Woman Suffrage Association (which Alice Paul's mother would later join).[2] This longstanding argument pushed the suffragists into making the right to vote their main focus.[12] The right to vote was now seen as the foundational right that would empower women to achieve every other reform.[13-15]

The shift in strategy from seeking federal change to pursuing state-by-state acceptance of suffrage presented its own challenges, primarily due to the political landscape at the time. By the late nineteenth century, the West (particularly Wyoming, Utah, and later Colorado) had enfranchised women, allowing them to vote, hold public office, and serve on juries. In contrast, the East and South continued with traditional politics and resisted reform. This regionalized difference presented the suffragists with a choice between continuing to grind through dozens of state-by-state campaigns or returning to the tactic of pursuing federal-level reform.[16]

Just as the political environment had changed over the decades that had passed since the initial Declaration of Sentiments had been announced, the social environment had also changed. By this time, there was also a determined, well-funded anti-suffrage movement.[17] They argued that the right to vote would coarsen female character, ruin the nuclear family, and destroy social

structure.[18] The suffragists not only had to convince the indifferent but overcome the actively hostile, which also contributed to the prolonged nature of their pursuit.

This was the political and social structure that Alice Paul, a completely different generation, inherited when she joined the suffrage movement. Born into a Quaker family, she was raised under the religious emphasis of the "Inner Light" existing in everyone regardless of gender.[2] As she later noted, "one of their principles was and is equality of the sexes. So I never had any other idea".[1] By the time she was a young girl, her mother was already involved with the suffrage movement, and Paul grew up attending meetings with her mother, allowing her to become familiar with its goals and methods.[3,19,20]

Another critical advantage Paul had in her era was her access to higher education. While Stanton remained bitterly indignant to the end of her life over her inability to attend Union College because of her gender, Paul suffered no such restriction. She attended Swarthmore College, which had been co-founded by her own grandfather.[21] There, she was taught by the leading female academics of the time, eventually going on to obtain graduate degrees in sociology and law.[1] Her education both empowered her with an in-depth knowledge of institutional power structure and the awareness that social change would remain ineffective until women obtained direct political power.

While pursuing her graduate studies in London, she joined the militant British suffrage movement started by Emmeline Pankhurst.[22] These radical suffragettes rejected the polite lobbying being used in America and instead pursued a "deeds, not words" approach to social change.[23] There, Paul experienced and learned the disruptive tactics she would later use in America. When Paul returned home to join the American suffrage movement, the goal was already established to focus strictly on the attainment of the right to vote. What she found, however, was a movement that had become stagnant in its progress, still trying to grind through a state-by-state approach that had been ongoing for decades.[24] The debate remained whether to continue with this localized method or to pivot back to the movement's original goal of federal-level reform. It was this strategic conflict, as well as institutional disagreement over the aggressive tactics that Paul proposed, that ultimately led her to break from the organization and create a completely different form of leadership.[22,25]

Two Leaders, Two Paths, One Movement

A fundamental difference in the leadership styles of Stanton and Paul arose from the profound difference in their ideologies. Stanton led the movement with a consistently broad focus. She wanted sweeping social reforms in all areas where she saw injustice for women—legal, economic, and even religious standing.[5] The vote to her was but one piece of a massive puzzle.[4] She campaigned in New York for property laws for women, which met with support in her movement, but also campaigned for liberal divorce laws, which was radical even among the reformers and was less accepted.

Her most controversial drive for reform was in what she regarded as the deepest source of women's insubordination: organized religion. Her volume *The Woman's Bible*, which

systematically analyzed the ways that the interpretation of scripture had been used to keep women inferior to men, was so controversial that her own movement formally published a rejection of it.[26] Her focus had gone beyond what the movement was able to support. Yet her broad ideology not only created the Women's Rights Movement's moral foundation but is the intellectual legacy which paved the way for later feminist movements.[27] Her leadership prioritized an intellectual framework over immediate legislative reform.[28]

Paul's vision for the movement, in contrast, remained on the singular goal of obtaining a federal suffrage amendment. In her own words, she states that she "never doubted that equal rights was the right direction. Most reforms, most problems are complicated. But to me, there is nothing complicated about ordinary equality".[29] Although in favor of reform in general, she rejected Stanton's global approach in favor of what had become over generations a focused goal. Stanton's broad approach had already previously fractured the movement due to her views on marriage and religion.[5] To have a unified front, Paul believed they needed to maintain a cause they all agreed upon. Adding other social issues such as birth control or labor laws risked once again fracturing their organization and thereby weakening it.[22] The right to vote was the only issue that various members of the coalition, which included women from all levels of socioeconomic backgrounds, were united on. Paul believed so strongly that they must remain united that when she formed the National Woman's Party (NWP), she banned them from pursuing any goal besides the federal suffrage amendment.[30] In her view, obtaining the right to vote was both attainable as well as the foundational change that would give women the power to enact future social reform.

Furthermore, Paul strongly believed that the most efficient and effective way to obtain the right to vote was through a federal constitutional amendment. The slow state-by-state approach that had been the tactic for several decades had only culminated in obtaining voting rights in a handful of states.[31] She felt that approach was inefficient, costing excessive amounts of time and money as they continued to petition all-male state legislatures. In addition, any freedom obtained by changing one state's constitution could be lost just by crossing state lines.[1] This view was supported by the history of the movement itself as generations of suffragists spent time and money and still were unable to make significant progress. Aiming for a solitary target—the White House—focused their energy where a single victory would create sweeping change.[24] As she put it, "when you put your hand to the plow, you can't put it down until you get to the end of the row".[32] That singular, unwavering focus was the hallmark of her approach. Only once the 19th amendment was finally ratified, did she pivot her goal to the attainment of equal rights, drafting the Equal Rights Amendment and devoting the rest of her life to pursuing its success.[33]

The starkest difference between the two leaders lies in their radically different strategies. Stanton relied on persuasion, in the form of writings and speeches. She drafted resolutions, delivered speeches to the legislature, and spent more than a decade on a lecture circuit carrying the argument for women's rights across the country. [8] Even while staying home with her growing family, she continued to draft speeches that Susan B. Anthony would travel across the country

delivering. As she famously put it, “I forged the thunderbolts and she fired them”. [6] Later in life, she helped write the *History of Woman Suffrage*, which shaped how the movement was viewed and remembered. [5] Ultimately, although her ideas were radical, her methods were the acceptable methods of reform in the nineteenth century: she pursued change by calm reasoning and rational appeal. [27]

Paul, on the other hand, used a far less conventional and significantly more disruptive route to change. During her time in England, she had engaged in the radical tactics of the British suffragette movement. [2] The experience of repeated arrests, hunger strikes, and forced feedings transformed her from a reserved social worker into a militant activist. Observing the impact of these methods, she noted that “the agitation has brought England out of her lethargy, and women of England are now talking of the time when they will vote, instead of the time when their children would vote”. [34] Consequently, she arrived in America with firsthand experience of the power of aggressive public exposure to capture media attention, and understanding exactly how public sympathy could be leveraged to create political pressure. Taking charge of the Congressional Committee of NAWSA, she launched her first large public demonstration—the 1913 Washington procession—the day before Woodrow Wilson’s inauguration. [1] Thousands of marchers descended on Washington, D.C., thrusting the suffrage movement onto the front page. When police failed to subdue the hostile crowds which attacked the marchers, the resulting scandal and subsequent congressional inquiry handed the movement a massive propaganda victory. [24]

Although initially supportive of Paul’s tactics, NAWSA’s leadership became increasingly uncomfortable with her radical approach. They worried that her aggressive public confrontations would label the movement “radical” and “unladylike,” thereby reversing the conservative political support they had earned through years of “respectable” lobbying. This tension led Paul to break from NAWSA and form the National Woman’s Party (NWP). [22] Through this platform, she launched her most militant campaigns. In 1917, she organized the “Silent Sentinels,” who became the first activists to picket the White House. [35] Their banners pointed directly at the hypocrisy of a wartime government defending democracy while simultaneously denying freedom to half of its own citizens.

When the picketers were arrested on charges of “obstructing traffic,” Paul escalated her resistance. She and her colleagues went on hunger strikes, endured brutal beatings, and suffered forced feeding. Authorities even subjected Paul to a psychiatric examination hoping to declare her insane—an effort that failed when the psychiatrist declared her courageous, not crazy. [2]

The most notorious abuse occurred at the Occoquan Workhouse in November 14,1917, a date Paul later dubbed the “Night of Terror”. [36] On that night, 33 jailed members of the NWP were brutally beaten and tortured. The media initially gave this event little notice -the country had just entered war, and the sentinels were partially viewed as unpatriotic due to their choice to picket the White House. With scant details beyond the arrest, there were more pressing issues occupying public attention. That narrative shifted radically due to Paul’s brilliant media tactic; she published firsthand accounts of the attacks in her newspaper, *The Suffragist*, and leaked

details to sympathetic journalists. [1] The graphic imagery of respectable women being smashed unconscious against iron bedframes, served worm-infested food, and having large tubes shoved into their noses to be force-fed was something even the staunchest opponents could not ignore. The brutal treatment of women holding American flags sparked a massive shift in public sympathy, making the administration's position increasingly untenable. [37]

Social scholars refer to this era as an example of the "radical flank effect". [38] For years, the more moderate NAWSA, under the leadership of Carrie Chapman Catt, had continued with conventional lobbying. Paul's radical approach, while increasing social pressure for change, also served to make Catt's mainstream methods look far more reasonable thereby increasing her access to negotiations with politicians. [24] The narrow single-vote victory in Tennessee demonstrated the degree to which increased social pressure was needed to bring this issue to the forefront of American politics. [39] Following years of political resistance, Woodrow Wilson finally announced his support for the federal suffrage amendment. [22] Paul had helped achieve victory through a steady tactic of relentless pressure rather than gentle persuasion. In doing so, she left a legacy of media-centered civil disobedience which would be used in later civil rights movements. [40]

Leadership: Two Eras, Two Leaders, One Movement

The shifting political and social landscape over the decades explains the arc of the suffrage movement and is critical to the analysis of its leadership. Elizabeth Cady Stanton's era faced the challenge of articulating systemic social injustices that nineteenth-century Americans failed to even recognize and then converting that recognition into organized support. [8] Stanton was the movement's foundation, framing the ideology that would bring "civilly dead" women into a world of gender equality. [5,41,42]

Conversely, Alice Paul's twentieth-century suffragists faced an entirely different challenge: turning passive sympathy into active support, overcoming the opposition of an organized anti-suffrage movement, and consolidating years of shifting goals and tactics into a focused approach to attain a definitive victory. [17,22] Stanton was the visionary who established the movement's ideological foundation; Paul was the strategist who converted those theories into fact. [3] Different eras facing different challenges required radically different leaders. [43]

Conclusion

Following the saga of the American women's suffrage movement is a profound lesson on the nature of leadership and change itself. The movement survived across multiple generations, evolving and ultimately succeeding due to the very differences in its leaders. These leaders were influenced by and adapted to the eras they found themselves in. Elizabeth Cady Stanton provided the foundation of the movement, naming its goals and creating an ideologic framework that would persist far beyond her lifetime. [5] Alice Paul launched off that foundation and the progress achieved in the decades before her, altering the movement's strategies to achieve victory. [22]

Analyzing the approaches of these very different leaders reveals that there is no single "right" strategy, but rather a strategy that is right for that movement at that specific moment in history. Stanton's persistent persuasion through writings and speeches were the acceptable social discourse of her time and allowed for a gradual growing sympathy for a cause that was originally deemed both ridiculous and dangerous. [8] Paul's militant approach, however radical, was launched when the movement was already accepted. Her tactics then flourished in a social atmosphere dominated by media exposure and a country defending democracy on the brink of war. [2] Either leader's approach executed at a different time in history would not have been possible. Stanton could not have achieved any level of acceptance for an inherently ridiculed cause using a militant approach, nor could Paul have continued with the sole tactic of ongoing discussion that had not achieved a final success despite over half a century of effort. [3] The final victory was a result of both types of leadership – leaders which established the framework for their cause, and leaders who understood what strategies were needed to orchestrate victory. [24]

From this perspective, the over seventy year-long suffragist movement was not prolonged due to the failure of earlier generations or defined by the success of the newer generation. It is the saga of change. It speaks about how slowly change occurs, how dedication must persist, and most importantly how their leaders must evolve. It requires leaders who have the vision to see the possibility of change, leaders who bring support to that vision, and leaders who craft the strategies to bring that vision into reality. And in this way, people change the world they live in.

Works Cited

1.Zahniser JD, Fry AR. Alice Paul: Claiming Power. New York: Oxford University Press; 2014.

2.Walton M. A Woman's Crusade: Alice Paul and the Battle for the Ballot. New York: Palgrave Macmillan; 2010.

3.Baker JH. Sisters: The Lives of America's Suffragists. New York: Hill and Wang; 2005.

4.Griffith E. In Her Own Right: The Life of Elizabeth Cady Stanton. New York: Oxford University Press; 1984.

5.Ginzberg LD. Elizabeth Cady Stanton: An American Life. New York: Hill and Wang; 2009.

6.Stanton EC. Eighty Years and More: Reminiscences 1815–1897. New York: European Publishing Company; 1898.

7.Boydston J. Home and Work: Housework, Wages, and the Ideology of Labor in the Early Republic. New York: Oxford University Press; 1990.

8.McMillen SG. Seneca Falls and the Origins of the Women's Rights Movement. New York: Oxford University Press; 2008.

9.Frederick Douglass on Women's Rights. Westport, CT: Greenwood Press; 1976.

10.Dudden FE. Fighting Chance: The Struggle over Woman Suffrage and Black Suffrage in Reconstruction America. New York: Oxford University Press; 2011.

11.DuBois EC. Feminism and Suffrage: The Emergence of an Independent Women's Movement in America, 1848–1869. Ithaca, NY: Cornell University Press; 1978.

12.Mead RJ. How the Vote Was Won: Woman Suffrage in the Western United States, 1868–1914. New York: New York University Press; 2004.

13.Lin A. The Fight for Women's Suffrage. 2022 [updated 2022; cited]; Available from: https://www.ohiovoterrights.org/post/the-fight-for-women-s-suffrage.

14.Hewitt N. Recasting Women's Votes: Race, Region and Rights 1869-1970. 2016. p. https://www.youtube.com/watch?v=VeeMVpKlBz8.

15.Sternberg A, Clark L. Fighting for the Vote: Women’s Suffrage in America. In: Challberg S, editor.2024. p. https://www.youtube.com/watch?v=71ERshnPxI8.

16.Kraditor AS. The Ideas of the Woman Suffrage Movement, 1890–1920. New York: Columbia University Press; 1965.

17.Marshall SE. Splintered Sisterhood: Gender and Class in the Campaign against Woman Suffrage. Madison, WI: University of Wisconsin Press; 1997.

18.Jablonsky TJ. The Home, Heaven, and Mother Party: Female Anti-Suffragists in the United States, 1868–1920. Brooklyn, NY: Carlson Publishing; 1994.

19.Marino K. Students, Suffrage, and Political Change: The College Equal Suffrage League and Campus Campaigns for Women’s Right to Vote, 1905–1920. The Journal of the Gilded Age and Progressive Era. 2021;20(3):370–91.

20.Schnell JC. Alice Paul. 2024 [updated 2024; cited]; Available from: https://www.ebsco.com/research-starters/history/alice-paul.

21.Swarthmore College. Pioneering Women. [Accessed August 9, 2026; cited]; Available from: https://www.swarthmore.edu/news-events/pioneering-women.

22.Lunardini CA. From Equal Suffrage to Equal Rights: Alice Paul and the National Woman's Party, 1910–1928. Lincoln, NE: University of Nebraska Press; 2002.
23.Pankhurst E. My Own Story. London: Eveleigh Nash; 1914.
24.Flexner E, DuBois EC. Century of Struggle: The Woman's Rights Movement in the United States. Enl. ed. ed. Cambridge, MA: Harvard University Press; 1996.
25.Magarey S. Dangerous Ideas: Women's Liberation. South Australia: University of Adelaide Press; 2014.
26.Kern K. Mrs. Stanton's Bible. Ithaca, NY: Cornell University Press; 2001.
27.DuBois EC. Woman Suffrage and Women's Rights. New York: New York University Press; 1998.
28.Michals D. Elizabeth Cady Stanton. National Women's History Museum; 2017 [updated 2017; cited 2026]; Available from.
29.Walton M. “Men and Women Shall Have Equal Rights”: Alice Paul and the Suffrage Movement. Gilder Lehrman Institute of American History; [Accessed August 9, 2026; cited]; Available from.
30.Library of Congress. Historical Overview of the National Woman's Party. [Accessed August 9, 2026; cited]; Available from.
31.Bill of Rights Institute. Alice Paul and the Struggle for Women's Suffrage. [Accessed August 9, 2026; cited]; Available from.
32.Fry AR. Conversations with Alice Paul: Woman Suffrage and the Equal Rights Amendment. In press 1976.
33.Library of Congress. From Nineteenth Amendment to ERA. [Accessed August 9, 2026; cited]; Available from.
34.Dinsmore DM. Alice Paul Returns from Great Britain—More Radical Than Ever. Ms Magazine. 2020.
35.National Park Service. Dr. Alice Paul. 2023 [updated 2023; cited]; Available from: https://www.nps.gov/people/alice-paul.htm.
36.National Park Service. Belmont-Paul Women's Equality National Monument. [Accessed August 9, 2026; cited]; Available from: https://www.nps.gov.
37.Ford LG. Iron-Jawed Angels: The Suffrage Militancy of the National Woman's Party, 1912–1920. Lanham, MD: University Press of America; 1991.
38.Haines HH. Black Radicals and the Civil Rights Mainstream, 1954–1970. Sociological Forum. 1987;2(4):649–53.
39.Weiss E. The Woman's Hour: The Great Fight to Win the Vote. New York: Viking; 2018.
40.Waging Nonviolence. Alice Paul's Enduring Legacy of Nonviolent Action. [Accessed August 9, 2026; cited]; Available from.
41.Yang M. An intimate dialog between race and gender at Women’s Suffrage Centennial. Humanities and Social Sciences Communications. 2020;7(1):65.
42.Marilley S, Judson SM. Woman suffrage & the origins of liberal feminism in the United States, 1820-1920. ProQuest; 1999 [updated 1999; cited 2026]; Available from: https://www.proquest.com/openview/846736f1a5f7cb552814610927c3fc01/1?pq-origsite=gscholar&cbl=30421.
43.Hogan LS. Elizabeth Cady Stanton, "The Solitude of Self". Voices of Democracy. 2014.

The Correlation of the Normalization of Mistreatment in Reproductive Health Systems and The Yellow Wallpaper By lillian Korzonkiewicz

Throughout history, women were expected to settle down and get married and have babies, and this was the prevalent ideal of a successful woman at the time "The Yellow Wallpaper" was published. The mistreatment and dismissive neglect shown to female patients by the reproductive industry is a societal issue, woven thread by thread throughout both history, and current day health systems. As we experience advancements in the medical world, a growing awareness of women's rights prompted women to reflectively look back on their personal experiences within the medical system. Many women of diverse socioeconomic and racial classes continue to experience various forms of disrespect, dismissal and blatant abuse by the system. This mistreatment comes in many forms, and can manifest as both verbal or physical abuse. Assumptions are made by physicians that an uneducated person has the inability to understand informed consent. This misnomer generates a dismissal of concerns that women may experience as a result of treatment.. This greatly affects a woman's mental and physical health, and when a doctor chooses to ignore a patient, this may influence a misdiagnosis. An ignored patient, if unable to address their concerns, could worsen from both a mental and physical standpoint. In "The Yellow Wallpaper," the narrator is confined to a room with bars on the windows, due to her diagnosis of postpartum hysteria and/or depression by her physician husband called John. This treatment isolates the narrator and eventually worsens her mental and physical health, as she becomes more and more fascinated by the unruly wallpaper in the room where she is kept. As her sanity spirals downwards, she starts to believe the wallpaper has a woman trapped inside of it. The narrator seeks to free the woman, corresponding with how she seeks freedom for herself. In "The Yellow Wallpaper", the narrator's mistreatment and overt disregard by her husband/doctor, reflect some of the current day practices which dismiss female patients from advocating for care in the reproductive health industry.

In the reproductive health field, many women experience mistreatment by male physicians, including gaslighting patients when they share concerns or question a diagnosis. In Understanding the Origins of a Social Catastrophe: Mistreatment in Childbirth as Normalized Organizational Deviance, Ramsey follows the common pattern of normalization. This normalization ranges from verbal abuse (ie, yelling, cussing, threatening) to physical or sexual abuse. Seventy percent of the women participating in this study experienced some form of abuse. The prevalence of abuse or neglect generated by doctors results in women accepting this treatment as a societal norm, throughout the healthcare system. The normalization not only marginalizes the severity of these experiences but also creates a dysfunctional cycle of compliance. There is a clear pattern of historic silence by women throughout medical observations and records. "The most reported behaviors were: being ignored or physically abandoned, in addition to the verbal and physical abuse. Observations of institutional deliveries in a rural setting found that nearly 70% experienced a form of mistreatment, indicating that

women also likely normalized mistreatment."(1) The common thread of mistreatment throughout these systems is a problem that continues to happen, even though it was common from the start of these institutions. In "The Yellow Wallpaper", Gilman addresses the narrator's hesitation to submit to the rest cure; as she is clearly uncomfortable, but her husband, as her doctor, believes that it is the best solution. The narrator believes she has no voice, and later in the story, as she tries to address her issues with the rest cure, her husband stifles her. He coddles her like a child as he does not let her voice her opinions, which results in dismissal with no consideration or discussion. "If a physician of high standing, and one's own husband, assures friends and relatives that there is really nothing the matter with one but temporary nervous depression - a slight hysterical tendency-what is one to do?(172)" the narrator states. She is unable to convey her opinion, but internally feels that something is wrong. She is silenced by the problems she is facing, as is a common occurrence for other women at the time.

In the reproductive health field, the violations of physical privacy and treatment practices closely correlate with the narrator's confinement in "The Yellow Wallpaper". The author struggled with postpartum depression and the narrator parrots her personal experiences with voicing her opinions on her treatment. It is common for many women, dissatisfied with their treatment in hospitals, to not voice their opinions or question protocol; similar to the normalization that is mentioned in the paragraph before. In "The Giving Voice to Mothers Study: Inequity and Mistreatment during Pregnancy and Childbirth in the United States.", Saraswathi mentions the issues that women have with time management with patients in a hospital setting. This generates problems for doctors who tend to spend less time and more lapses in the care and monitoring of new mothers.

This problem highlights the failure to address the urgency required to ensure safety and healing for postpartum women. Many women also addressed the problems of violation in their personal privacy. This violation can create an unsafe environment for women. There were also women who reported issues of their healthcare provider withholding necessary treatment. This led to their issues worsening, and patients losing trust in their doctors and the overall industry when "requests for help in a reasonable amount of time" (7.8%) were ignored or delayed. Fewer women reported overt violation of physical privacy (5.5%), and healthcare practitioners threatening to withhold treatment or forcing them to accept treatment they did not want (4.5%)" (3).

The neglect these patients experience with their care is closely related to the mimic the author's unquestioning submission and experiences in The Yellow Wallpaper. The narrator states; " I take phosphates or phosphites, whichever it is, and tonics, and journeys, and air, and exercise, and am absolutely forbidden to "work" until I am well again. Personally, I disagree with their ideas. Personally, I believe that congenial work, with excitement and change, would do me good. But what is one to do?"(172) The Narrator does not consent to this treatment, as her husband/doctor forces her into the rest cure. She is given medication, ignorant of both the prescription name and potential side effects of the treatment meds. The narrator is provided no control over her bodily anatomy, and this is similar to some of the women in this study. Her

doctor / husband disregards her own beliefs about what would benefit her. This illustrates a problem with the communication between her husband and the lack of respect from him for her own agency. Both of these quotes portray a disturbing pattern where women's voices and preferences are sidelined in favor of the common medical practices of the time. As the narrator in “The Yellow Wallpaper" insists that she does not agree with their ideas but has no choice. It reflects a broader sentiment among women today and back in the time of “The Yellow Wallpaper", who feel constrained by their health care providers who impose treatments on them that they neither asked for, nor wanted.

In both “The Yellow Wallpaper" and current day society women face systematic barriers accessing healthcare, which is demonstrated by the disregard the narrator faces herself in the industry with her doctor/ husband. The article“In Collective Persistence, Care, and Advocacy amid Repeated Attacks on Reproductive Freedom.”, The *American Journal of Public Health* highlights the issues women must navigate in the reproductive system, when facing a costly and at times uninsurable access to healthcare. Women who may not be able to care for a child often seek abortion, which may be an extremely costly and unsafe endeavor due to the risks people may face, caused by issues of legality in certain states. These disenfranchised women may have to seek a non-legal abortion in their state if they can't travel to a more permissive one. Unable or unsure of how to advocate for themselves, women are not even able to be able to support themselves, and certainly not a child so they seek abortion. The article states:

> She must navigate systematic, costly, and at-times-insurmountable barriers to access, as well as uncertainty about how, where, and when they will get care. In turn, they bear psychological and health outcomes of state-imposed care delay and denial. Providers cannot employ the full scope of their practice without facing legal repercussions, an emotional burden that can affect their health and well-being.

This reinforces the theme of inaccessibility these women have in the medical reproductive field. In the yellow wallpaper the narrator states; **“**John says if I don't pick up faster he shall send me to Weir Mitchell in the fall. But I don't want to go there at all” (177).
The narrator is reluctant to receive treatment but is not given a choice due to her husband/ doctor's authority. John has complete control over her medical autonomy due to him being both her doctor and her husband. Both of these quotes illustrate a frustrating reality; they illustrate the systematic problem that women are often trapped in medical organizations that prioritize institutional protocols over their patients' own individual needs and preferences. The anxiety that the narrator experiences connects to the idea of the problems of access with reproductive health as a postpartum woman. As the narrator received terrible treatment but was worried about receiving even worse treatment in another area away from her family.

In conclusion “The Yellow Wallpaper” serves as a powerful lens Through the ongoing mistreatment in reproductive Medical Systems back in the 1890’s to current day. “The Yellow

Wallpaper" reveals the narrator's experiences of isolation and dismissal reflecting the systematic barriers many women face today. As mistreatment is common in the reproductive health industry parallels are drawn through the narrator's struggle and the realities of female patients, which highlights the persuasive culture of normalization, which silences women's voices within the healthcare system. The normalization of mistreatment in both the medical systems of the current day and the other wallpaper display the urgent and ongoing need for a systematic reform. The cycle of neglect and abuse in reproductive healthcare systems can only be changed through confrontation of these problems. It is essential that we remember the past and continue to evolve in reproductive health treatments, so women aren't put into uncomfortable and unnecessary situations due to the social standpoints affecting them.

Works Cited

K., Ramsey. “Understanding the Origins of a Social Catastrophe: Mistreatment in Childbirth as Normalized Organizational Deviance .” Proquest.com, 2024, https://academiccommons.columbia.edu/doi/10.7916/d5gp-2z11 Accessed 2 June 2025.

Rice, W. S.,DrP.H.M.P.H., Hartwig, S. A., M.P.H., McMahon, H. V.,M.S.P.H.C.P.H., Irwin, C. R. N. P., Quinones, N., M.P.H., Redd, S. K.,PhD.M.S.P.H., & Singh, K. A., M.S.P.H. “Collective Persistence, Care, and Advocacy amid Repeated Attacks on Reproductive Freedom.” Proquest.com, American Journal of Public Health, 2024, www.proquest.com/scholarly-journals/collective-persistence-care-advocacy-amid/docvi. Accessed 2 June 2025.

Vedam, Saraswathi, et al. “The Giving Voice to Mothers Study: Inequity and Mistreatment during Pregnancy and Childbirth in the United States.” Reproductive Health, vol. 16, no. 1, June 2019, https://doi.org/10.1186/s12978-019-0729-2. Accessed 2 June 2025.

The Flaws of the Indian Reorganization Act of 1934 By Solomon Heyman

The Indian Reorganization Act of 1934 sought to improve conditions on Native American Reservation, and to assuage previous failures in tribal policy making. Decades before the Act's passage, a different piece of legislation called the Dawes Act, or the General Allotment Act, had divided Native lands into family portions in order to effectively "...turn the Indian into a homesteader" (Taylor 4). In practical terms, this Act aimed to impel Native Americans to acquire Americanized traits, rather than continue living on their lands in collective societies. But Americanization was just one consequence of the law. Excess land that resulted from creating allotments was confiscated by the United States, further limiting the territory controlled by Natives. Between 1887 and 1932 Native-controlled land diminished dramatically from 139 million acres to 48 million acres, significantly reducing Indian sovereignty (Taylor 5). During the New Deal, John Collier, head of the Bureau of Indian affairs (BIA) sought to institute the Indian Reorganization Act (IRA) of 1934: an Act that would end allotment activities and reinforce Indian sovereignty through self-governing bodies on respective reservations (Taylor 17, 28). Although the IRA marked a significant transformation from the philosophy that shaped the Dawes Act of 1887, with respect to physical Indian autonomy there was no significant change. This raises a critical question: did the Indian Reorganization Act fail to solve the Indians' problems or was it a crucial step in Native societal progression? Historians like Graham Taylor and Richmond Clow agree that the Indian Reorganization Act was flawed, but disagree about the nature of its failure. Taylor argues that the Act was flawed due to budgetary constraints by Congress and legal definitions that gave the term "tribe" too broad a meaning. Clow, on the other hand, suggests that the Act was entirely unnecessary, as Native groups had functioning governments already. Ultimately, three flaws proved to be crucial: under the Indian Reorganization Act of 1934, the BIA retained too much power, Native perspectives were not considered when drafting the Act, and the system of governance implemented on reservations reflected U.S. democratic norms, rather than traditional native values.

The Dawes Act of 1887 marked a crisis for both the lives and outlooks of Native Americans. The Act, sponsored by Senator Henry Dawes, aimed at procuring Native lands and redistributing the territories as small allotments to individual Native families. Essentially, the U.S. government sought to dismantle the communal aspect of Native American life and assimilate Natives into mainstream American culture. Many Native Americans regarded these moves, and particularly the emphasis on religious assimilation, as a significant threat to their traditions and way of life (Taylor 3). Further compounding the Americanization that was catalyzed by the Dawes Act, any excess land that was not appropriated as allotments was seized and sold to white settlers. By numbers, the policies instituted through the Dawes Act proved disastrous: according to one study, more than half of Native Americans living in the United States in 1928 earned under two hundred dollars each year, compared to the average American family's two thousand; fewer than five percent made more than five hundred (Taylor 7). Rather than successfully assimilating Natives into American society, the Dawes Act did the opposite: it

essentially excommunicated Native Americans, limiting their exposure to crucial markets and as a result causing widespread poverty. Ultimately, the Dawes Act stood as a major hindrance in the societal progression of Native Americans within the United States.

The Meriam Report, published shortly before the IRA and formerly known as the "The Problem of Indian Administration," was commissioned by the Institute for Government Research and headed by statistician Lewis Meriam. Its primary purpose stemmed from suspicions that significant flaws in the Dawes Act had greatly affected Native American life. The findings of the document delineated the dire life of Indians confronted with abhorrent living conditions and health issues such as widespread tuberculosis. The Meriam Report noted that, "The housing conditions are likewise conducive to bad health. Both in the primitive dwellings and in the majority of more or less permanent homes which in some cases have replaced them, there is great overcrowding, so that all members of the family are exposed to any disease that develops, and it is virtually impossible in any way even partially to isolate a person suffering from a communicable disease" (Meriam 4). These findings confirmed that the Dawes Act had detrimental effects on Natives, and that more legislation was called for in order to reverse these impacts and improve conditions for Natives.

The Indian Reorganization Act of 1934, also known as the Wheeler-Howard Act, was headed by John Collier, a prominent voice for reform in federal-Tribal affairs (Taylor 12). Appointed by Franklin D. Roosevelt, Collier was tasked with restoring Indigenous welfare after the Meriam Report which revealed the many tribulations the Natives faced. The first goal was to restore tribal sovereignty by requiring tribes to write their own constitutions and maintain their own governments on reservations. Another crucial facet of the Act was putting an end to land allotments, and allowing tribes to collectively own land. To negate the massive reduction in land following the Dawes Act of 1887, governments bought back some land from white settlers and returned it to Native Americans. To protect tribal culture, the government ended their attempt to Americanize the Natives and instead barred the BIA from interfering with Indigenous religious practices. This Act, with sole aims at achieving Indian prosperity and livelihood, marked a crucial turning point in the treatment of the Natives in the United States, and attempted to pave the way for stability on tribal reservations.

Graham D. Taylor, the historian and author behind the book *The New Deal and American Indian Tribalism,* a study that examines the effectiveness of the IRA, posits that the act was inherently flawed and failed to solve many of the Indians' issues. Taylor argues that "...by making the transfer of lands from individual to tribal ownership voluntary…[the Act] provided only a limited degree of control to the federal agency in reshaping Indian economic resources" (Taylor 31). Taylor also argues that the IRA did not consider what a "tribe" really meant. He expresses this idea by quoting a line in the last section of the act: "any Indian Tribe, organized band, pueblo, or the Indians residing on one reservation" (Taylor 63). He argues that the BIA attributed too vague of a definition to the term "tribe," which meant that some groups lacked sovereignty under the federal government and could not access federal resources. Taylor also claims that the BIA did not allocate enough autonomy to tribes when they drafted their own

constitutions. Rather than allowing the tribes to draft their own constitutions and self-govern, which was a focal point of the IRA, the BIA gave them model constitutions from which to create their own (Taylor 97). Taylor also notes that implementation of the IRA was hampered by budget shortfalls for the BIA and excessive congressional oversight: "Regardless of their initial commitment to genuine Indian autonomy, administrators felt constrained by budgetary limitations and the growing intensity of congressional criticism to channel their main efforts into economic rehabilitation of the reservations" (Taylor 150). Rather than focusing on long-term solutions for stability, the BIA instituted short-term economic recovery programs that were cheaper than finding a lasting solution. Ultimately, Taylor provides concrete evidence as to the flawed nature of the IRA, and utilizes primary sources as means to support his opinions.

Based on studying two Oglala Sioux reservations, Rosebud and Pine Ridge, historian Richmond L. Clow argues that the IRA had two major flaws. First, the Sioux already had a functioning government, and did not need American regulators to teach them how to live democratically. Second, the IRA's version of "self-government" was misleading: the constitutions that the IRA created limited tribal autonomy, and created a governance model in which U.S. politicians executed the tribes' affairs. In order to corroborate his first argument that the Sioux already had a functioning government, Clow notes that "The Brule' Sioux…and their Oglala kinsfolk on the Pine Ridge had between them written and adopted seven constitutions between 1916 and 1933" (Clow 125). BIA head John Collier's call to "set the feet of our Indian friends on the path that leads to self-government" aside, Clow insists that tribes "already possessed the powers of self-government" (Clow 131). If anything, Clow argues, Collier's interventionist approach prevented tribes from achieving "self-government" by heavily restricting their autonomy.

One major flaw that plagued the Indian Reorganization Act from its inception was the Bureau of Indian Affairs' tendency to withhold power from being allocated to tribes. This pattern was extremely evident during the collaborative interactions between the Bureau and tribal governments. The IRA's provisions clearly delineate this imbalanced relationship: "The Secretary of the Interior, if he shall and it to be in the public interest, is hereby authorized to restore to tribal ownership…" (IRA Sec. 3). Through the wording of the statement, "if he shall and it to be in the public interest," the Act places authority in the hands of a government official. Instead of valuing the welfare and needs of the tribe, the BIA did what was best for the overall public. Another example of text in the Act that implies a restriction of power is, "damages shall be paid… as may be determined by the Secretary of the Interior" (IRA, Preamble). This line clearly vests in the Secretary, and not Tribal leaders, the power to decide which groups are compensated and with how much. This uneven distribution of power undermined the goal of tribal sovereignty, and skewed the decision-making towards the aims of the U.S. government.

Further complicating matters, Native perspectives were not considered when drafting the IRA. In a 1943 letter addressed to Native Americans, Secretary of the Interior and IRA head John Collier expressed his concerns regarding precisely this flaw in the federal government's operations. In the second section of the letter, Collier proposes a change in methodology as to

how joint policy making should be performed: “Plans for Indian advancement must be not only accepted but cooperatively devised by the Indians themselves” (Collier 4). Here Collier admits to having previously created policies for the Native Americans rather than give them the autonomy to have the leading say in legislation. His need to clarify this distinction reveals that the Bureau for Indian Affairs held too much power, and could supersede the authority of Tribal leaders, a non-optimal arrangement if the primary objective is to establish Indian self-governance. Collier yet again hints at this lopsided relationship in policymaking when he says during a final remark in his letter: “These programs will be subject to approval by this office, and they will be subject to change and growth as we and the Indians come to grips with their further refinement and execution” (Collier 6). Through the clarification that the previously enumerated programs “... will be subject to approval by this office,” Collier indirectly shows that policies are being heavily skewed towards the BIA’s wants and not the perspectives of the Indian tribes. Although the IRA sought to prioritize the perspectives of the Indians and put full effort into championing their opinions, Collier’s diction makes clear that they should have been included in the process of designing the law.

Rather than maintaining a traditional Indian government that was ruled mainly by a chieftain, the BIA instituted policies of US democratic governance. The Indian Reorganization Act, despite its intentions of constructing a viable governance model for reservations, swayed too far from the traditional Indian tribal systems of leadership. One major defect regarding the reforms implemented by the BIA was imposing an American-style government rather than a model that resonated with traditional tribal administration. Lewis Morgan, an author who wrote about the Iroquois Confederacy in 1851, provides illuminating context about the tribal structure of native nationhood. In his book *League of the Ho-De-No-Sau-Nee*, Morgan writes: “In the case of the Ho-De-No-Sau-Nee, the organization was externally so obscure, as to induce a universal belief that the relations between ruler and people were simply those of chief and follower”(Morgan 62). Here, Morgan conveys that the Ho-De-No-Sau-Nee governments were merely composed of chiefs and their followers. He then writes: “It thus appears that the government of the Iroquois was an oligarchy, taking the term, at least, in the literal sense, ‘the rule of the few’” (Morgan 63). The Iroquois, which is another name for the Ho-De-No-Sau-Nee Confederacy, according to Morgan, maintained an oligarchical government in which the chief primarily led. This backdrop of an oligarchical government contrasts heavily with the government instituted under the Indian Reorganization Act. In John Collier’s letter, written during the implementation of the IRA and directed towards the superintendents of the tribal councils, he writes: “‘planning’ must be democratic and participative, not authoritative and directed” (Collier 4). Collier urges superintendents to implement a “democratic and participative” governance model, which resembles an American style of governance. Collier goes on to claim that U.S. “officials” have too much power and influence in tribal policy-making. This distinction between the pre-IRA oligarchachal form of government conveyed through the Iroquois and the post-IRA government which is expanded upon in Collier's letter, shows the drastic effect the Act had on tribal governance.

The Indian Reorganization Act of 1934 reflected a significant shift in tribal government and life for the indigenous peoples. While it marked a notable change from the 1887 Dawes Act, in both ideology and legislation, three crucial details proved to hinder the Act's success. First, the Bureau for Indian Affairs retained too much control over Indian self-governance. Second, the Act was drafted without considering Native points of view or needs. Third, the Act insisted that reservations govern themselves according to U.S. democratic norms, a drastic shift from the 19th-century traditional structure of most tribes. While today's tribes have self-governance and control of their economies, many still struggle with high rates of poverty and inter-tribal conflict, the long-term effects of failed nineteenth- and twentieth-century legislation.

Ultimately, while the IRA failed to fully restore tribal sovereignty, it marked a critical step in advancing U.S.-Tribal affairs, and serves as a precedent for further progress.

Works Cited

Clow, Richmond L. “The Indian Reorganization Act and the Loss of Tribal Sovereignty: Constitutions on the Rosebud and Pine Ridge Reservations.” Great Plains Quarterly, vol. 7, no. 2, 1987, pp. 125–34. JSTOR, http://www.jstor.org/stable/23530981. Accessed 8 May 2026.

Collier, John. “Circular No. 3537 (‘To: Superintendents, Tribal Councils, All Indian Service Personnel and All Indians’).” Nov. 15, 1943. Records of the Bureau of Indian Affairs, Record Group 75, Correspondence. National Archives Records Administration, Riverside. https://catalog.archives.gov/id/608241614?objectPage=8. Accessed 7 May 2026.

Indian Reorganization Act. “An Act of June 18, 1934, Public Law 73-383, 48 STAT 984, to Conserve and Develop Indian Lands and Resources; To Extend to Indians the Right to Form Business and Other Organizations; To Establish a Credit System for Indians; To Grant Certain Rights of Home Rule to Indians; To Provide for Vocational Education for Indians; and for Other Purposes.” National Archives and Records Administration, National Archives and Records Administration, 18 June 1934, catalog.archives.gov/id/7873515.

Meriam, Lewis. “The Problem of Indian Administration.” 1928. Internet Archive, archive.org/details/in.ernet.dli.2015.276054/ page/n23/mode/2up. Accessed 7 May 2026.

Morgan, Lewis Henry. “League of the Ho-Dé-No-Sau-Nee, or Iroquois.” Internet Archive, Rochester : Sage & Brother ; New York : M.H. Newman & Co. ; [etc.], 1 Jan. 1970, archive.org/details/leagueofhodnos00inmorg/page/62/mode/2up.

Taylor, Graham D. The New-Deal and American Indian Tribalism the Administration of the Indian Reorganization Act, 1934-45 by Graham D. Taylor. Univ. of Nebraska Press, 2015.

Precarity or Belonging? Job Insecurity, Social Capital, and Well-Being among Immigrant Workers in the Chinese Ethnic Enclave Economy By Elaine Liu

Abstract

Decades of research on the immigrant enclave economy have asked whether employment in co-ethnic businesses helps or harms the workers who depend on it, and that debate has been conducted almost entirely in the currency of wages and labor compensation. This study asks a different question. Drawing on survey data from 179 Chinese-speaking immigrant workers employed in ethnic Chinese businesses in the greater Los Angeles area in 2026, it first maps what these workers themselves identify as the sources of their job insecurity, and then examines how perceived job insecurity, co-ethnic bonding social capital, and cross-ethnic bridging social capital relate to their financial and psychological well-being. Workers located the threat in the market rather than the state, ranking the broader economic environment, competition from other workers, and displacement by lower-wage labor as the strongest sources of insecurity, while immigration status ranked lowest of the eleven sources assessed. Perceived job insecurity was the dominant negative correlate of both financial and psychological well-being. Bonding social capital, contrary to hypothesis, predicted neither outcome. Bridging social capital and English proficiency were positively associated with both.

Introduction

A significant number of the immigrant workforce in the United States is employed not in the mainstream labor market but in businesses owned and operated by co-ethnics. In 1980, Kenneth Wilson and Alejandro Portes termed this sector the immigrant enclave economy, defining it by the concentration of immigrant-owned firms, their orientation toward co-ethnic or general markets, and their employment of a significant proportion of the co-ethnic labor force (Wilson and Portes). For immigrants whose access to employment is limited by language barriers or discrimination, enclave businesses are often the most accessible point of entry into paid work, and for the Chinese immigrant communities of greater Los Angeles, one of the largest and longest-established ethnic economies in the country, they remain a central institution of working life (Light and Gold; Zhou, *Chinatown*).

Whether enclave employment benefits the workers who depend on it has been debated for a long time. Wilson and Portes argued that the enclave constitutes a distinct mode of labor market incorporation in which co-ethnic employment yields returns unavailable to immigrants, and that enclave businesses function as entrepreneurial incubators. Sanders and Nee argued that the economic benefits of the enclave accrue primarily to owners, whereas employees earn lower returns on their human capital than comparable workers in the open economy. Subsequent research has complicated rather than resolved the disagreement: enclave returns appear to vary by gender (Zhou and Logan), by cohort and mode of incorporation (Portes and Shafer), and by how the enclave itself is operationalized (Portes and Jensen, “The Enclave and the Entrants”; Xie and Gough). Ethnographic work has meanwhile documented conditions of enclave employment

that formal earnings comparisons do not capture, including thin margins, informal employment relations, and labor discipline enforced through community obligation (Kwong, *The New Chinatown*; Kwong, *Forbidden Workers*; Chin; Hondagneu-Sotelo).

Two limitations of this literature motivate the current study. First, the debate has been conducted almost entirely over wages and returns to human capital. Financial security and psychological well-being, meanwhile, have not received enough research attention (Diener et al.; Netemeyer et al.). Second, current literature has been better at characterizing the enclave's structural conditions than asking workers themselves what makes their position feel insecure. Contemporary enclave workers face a more precarious environment than their predecessors, including automation and artificial intelligence, shifting immigration policy, and there is little systematic evidence on how workers rank these threats.

This study seeks to address both limitations with original survey data from Chinese-speaking immigrant workers employed in ethnic Chinese businesses in the greater Los Angeles area. It first maps the perceived sources of these workers' job insecurity, and then tests how perceived job insecurity and co-ethnic bonding and bridging social capital relate to two dimensions of well-being, financial and psychological.

Greater Los Angeles is an appropriate site for this question. The Los Angeles metropolitan area is home to roughly 630,000 residents of Chinese origin, second only to New York (Pew Research Center), and its Chinese ethnic economy is unusual in both form and scale. Since 1970s, Chinese settlement moved outward from the downtown Chinatown into the San Gabriel Valley, where Monterey Park became the first city in the continental United States with a majority Asian American population and the anchor of what Li called an ethnoburb, a suburban cluster of co-ethnic residence and commerce (Li). The resulting ethnic economy now extends across more than twenty San Gabriel Valley municipalities and spans restaurants, supermarkets, retail and professional offices, so that a Chinese-speaking immigrant can enter this labor market, advance within it, and remain in it with little sustained contact with English-language employers. Two features of the setting also make it analytically useful. Its industrial breadth means that enclave employment here is not confined to the garment or restaurant sectors, which made the focus of earlier Chinatown studies, and its maturity means that workers are embedded in a dense and long-established co-ethnic institutional environment rather than an emerging one.

Literature Review and Hypotheses

Job Insecurity in Enclave Employment

Perceived job insecurity refers to the subjectively experienced threat of involuntary job loss and the worry that accompanies it (De Witte; Vander Elst, De Witte, and De Cuyper). It is conceptually distinct from actual job loss, but its consequences can be severe: research studies find robust negative associations between perceived job insecurity and psychological health, job attitudes, and life satisfaction (Sverke, Hellgren, and Näswall; Cheng and Chan). The stressor

can also operate through anticipated rather than realized loss, which is why perception is the appropriate level of measurement.

Enclave employment concentrates the structural conditions that produce such perceptions. Enclave businesses compete in saturated co-ethnic niches with thin margins and high labor substitutability (Sanders and Nee; Zhou, *Chinatown*; Aldrich and Waldinger), employment relations are frequently informal and lack contractual protection (Kwong, *The New Chinatown*), and many workers hold human capital that is oriented toward the enclave rather than recognized by the mainstream economy (Xie and Gough). To these long-standing conditions, recent years have added new ones: anxiety about displacement by artificial intelligence and automation, and uncertainty produced by shifts in immigration policy. To explore and identify which factors contribute to enclave workers' sense of job insecurity, I propose the following research question:

> **RQ1:** Which sources of job insecurity do immigrant workers in ethnic enclave businesses perceive as most influential?

Whatever its sources, perceived job insecurity should erode well-being through two distinct channels. Financially, insecure employment makes future income unpredictable, undermining the sense of security and control over one's financial life that defines financial well-being (Consumer Financial Protection Bureau). Psychologically, the anticipation of loss functions as a chronic stressor associated with strain, impaired functioning, and distress (De Witte; Sverke, Hellgren, and Näswall). Accordingly, I propose the following two hypotheses:

> **H1:** Perceived job insecurity is negatively associated with financial well-being.
> **H2:** Perceived job insecurity is negatively associated with psychological well-being.

Bonding Social Capital in the Co-Ethnic Community

Social capital, the resources embedded in social networks, takes different forms depending on network structure. Putnam distinguishes bonding social capital, the trust, reciprocity, and mutual aid within dense, homogeneous groups, from bridging social capital, the connections that span group boundaries. Ethnic enclaves are, almost by definition, rich in the bonding form: the same closure that concentrates co-ethnic firms and workers also generates enforceable trust, mutual assistance, and privileged information flows within the community (Portes and Sensenbrenner; Sanders and Nee).

The consistent findings from the social support literature suggest that such resources protect well-being. Perceived support buffers the impact of stressors on psychological outcomes (Cohen and Wills), and community ties supply both instrumental aids, including emergency loans and job information, and the sense of belonging associated with better psychological functioning (Keyes). To formally test these hypotheses, I propose:

H3: Bonding social capital is positively associated with financial well-being.
H4: Bonding social capital is positively associated with psychological well-being.

Bridging Social Capital and the Mainstream Opportunity Structure

If bonding capital is the form the enclave generates in abundance, bridging capital is the form it tends to make scarce, and the two are theorized to do different work according to existing social capital literature. According to Briggs and Putnam, bonding capital is suited to getting by while bridging capital is suited to getting ahead: within-group ties circulate the resources a community already possesses, whereas ties that span group boundaries supply access to new information, unfamiliar opportunities, and mainstream institutions. Granovetter's classic work on the strength of weak ties also supports this prediction: non-redundant information travels along the connections that reach outside a dense cluster, because everyone inside the cluster already knows the insider information. For enclave workers, this is likely to apply because when the co-ethnic labor market is saturated, the employment and institutional access that would improve a worker's social position lie largely outside the group and are reached, if at all, through ties that cross its boundary.

The limited direct evidence is consistent with this reasoning. Lancee found that immigrants' economic outcomes in the Netherlands are significantly contributed by bridging but not to bonding capital. In terms of predicting one's psychological well-being, cross-group ties may reduce isolation and broaden the sense of belonging beyond a single community, but contact across ethnic boundaries can also expose immigrant workers to discrimination and to the strain of navigating unfamiliar institutional settings, so the net association is not obvious in advance.

Given the uneven strength of the theoretical expectations across the two outcomes, it is posed as a research question rather than a directional hypothesis:

RQ2: How is bridging social capital associated with the financial and psychological well-being of immigrant workers in ethnic enclave businesses?

Well-Being as the Outcome, and the Role of Acculturation-Related Resources

Consistent with the multidimensional view of well-being, the study measures two conceptually distinct outcomes. Financial well-being is the perceived ability to meet obligations, absorb shocks, and feel secure about one's financial future (Consumer Financial Protection Bureau). Psychological well-being is operationalized as recent positive functioning and absence of distress, following the General Health Questionnaire tradition (Goldberg and Williams).

Finally, because the sample comprises first-generation immigrants, acculturation-related resources, particularly English proficiency and length of U.S. residence, are theoretically relevant to both outcomes. Specifically, language proficiency conditions access to mainstream labor markets, institutions, and information (Berry; Xie and Gough), and is therefore entered, along with demographic characteristics, as a control variable in all models, with its independent associations reported.

Method

Sampling and Participants

Participants were recruited through the author's personal connections and convenience sampling. To be eligible, individuals had to be 18 years of age or older and to be currently working, or to have previously worked, in an organization located in the greater Los Angeles area, with the additional requirement that the employing business operate as an ethnic Chinese business. Data were collected through an online, self-administered Chinese-language questionnaire hosted on Jotform in two waves between April and June 2026 (Wave 1 $n = 52$; Wave 2 $n = 127$), yielding a combined analytic sample of 179 Chinese-speaking immigrant workers. Independent-samples comparisons indicated that the two waves did not differ substantively on the focal study constructs, and the waves were therefore pooled together for analysis.

Slightly over half of the sample were female (55.9%; 43.0% male, 1.1% preferred not to report gender). The sample was relatively mature, with 55.9% aged 45 or older (31.8% aged 55 or older), and most participants were married (63.7%). The great majority traced their origin to Mainland China (82.1%), followed by Taiwan (11.2%), and most had resided in the United States for more than ten years (54.2%). Educational attainment was relatively high: 48.6% held at least a bachelor's degree, with an additional 29.1% having completed an associate or vocational program. English proficiency was mixed: 36.9% rated their English as very good, whereas 31.3% rated it as not very good or reported no English, and most respondents worked bilingually (45.8%) or primarily in Chinese (22.3%). Just over half were employed in professional, technical, office, or finance occupations (51.4%), with the remainder concentrated in construction and manual labor (19.6%), personal services such as nail, massage, and hair care (15.1%), and food service (9.5%). Most worked full-time (80.4%), and 37.4% were self-employed business owners. Monthly household income spanned all bands, with 34.6% reporting $8,000 or more.

Measures

Sources of Job Insecurity. To address RQ1, participants rated the extent to which each of eleven potential sources contributed to their job insecurity on five-point scales (1 = *no effect at all*, 5 = *a very strong effect*). Sources included the current economic environment, competition from other workers, being replaced by lower-wage workers, artificial intelligence replacing jobs, English ability, age or health, lack of U.S.-recognized transferable credentials, changes in immigration policy, lack of a formal labor contract, employer business viability, and immigration status. These items were treated as descriptive indicators and were not combined into a composite.

Perceived Job Insecurity. Perceived job insecurity was measured with four items adapted from the Job Insecurity Scale (De Witte; Vander Elst, De Witte, and De Cuyper). Participants rated their agreement with statements such as "I think I may lose my job in the near future" and

"I feel uneasy about the future of this job" on five-point scales (1 = *strongly disagree*, 5 = *strongly agree*). One positively worded item ("I am confident I can keep my current job") was reverse-coded so that higher scores denote greater perceived insecurity ($M = 2.40$, $SD = 0.75$, Cronbach's $\alpha = .91$).

Bonding Social Capital. Bonding social capital was assessed with six items capturing trust, mutual aid, belonging, and information sharing within the co-ethnic Chinese community (e.g., "When I have problems, I can turn to other members of the Chinese community for help"; "People in the Chinese community look after one another"), rated on five-point agree-disagree scales ($M = 3.73$, $SD = 0.54$, Cronbach's $\alpha = .95$).

Bridging Social Capital. Bridging social capital was assessed with six items measuring connections across ethnic and cultural boundaries (e.g., "I frequently interact with people of other ethnic or cultural backgrounds"; "I feel accepted by people of other cultural backgrounds in my area"), using the same response format ($M = 2.97$, $SD = 0.90$, Cronbach's $\alpha = .97$).

Financial Well-Being. Financial well-being was measured with five items adapted from the Consumer Financial Protection Bureau Financial Well-Being Scale, tapping perceived financial security and control (e.g., "I am just getting by financially"; "I have money left over at the end of the month"), rated on five-point scales (1 = *does not describe me at all*, 5 = *describes me completely*). The four negatively worded items were reverse-coded so that higher scores reflect greater financial well-being ($M = 3.03$, $SD = 0.83$, Cronbach's $\alpha = .89$).

Psychological Well-Being. Psychological well-being was measured with eleven items adapted from the twelve-item General Health Questionnaire (GHQ-12; Goldberg and Williams), assessing recent positive functioning and the absence of distress (e.g., "Been able to concentrate on what you are doing"; "Felt constantly under strain"), rated on five-point frequency scales. The six distress items were reverse-coded so that higher scores indicate better psychological well-being ($M = 3.51$, $SD = 0.47$, Cronbach's $\alpha = .87$).

Control Variables. Demographic and background variables, including age, gender, educational attainment, years of U.S. residence, English proficiency, and monthly household income, were included as covariates.

Analysis Strategy

To test the hypothesized associations of perceived job insecurity (H1, H2) and bonding social capital (H3, H4) with the well-being outcomes, a series of hierarchical multiple regression models was estimated, with financial well-being and psychological well-being as the dependent variables. In each model, demographic covariates (age, gender, education, years in the United States, English proficiency, and household income) were entered in the first step, and the focal predictors, perceived job insecurity and bonding social capital, were entered in the second step; categorical predictors were dummy-coded. This sequencing isolates the incremental variance explained by the focal predictors beyond demographic controls.

RQ1 was addressed descriptively by rank-ordering the mean perceived influence of the eleven sources of job insecurity. Descriptive statistics and scale reliability estimates for all study variables appear in Table 1.

Table 1

Descriptive Statistics of Major Study Variables (N = 179)

Variable	M	SD	α	No. of items
Perceived job insecurity	2.40	0.75	.91	4
Bonding social capital	3.73	0.54	.95	6
Bridging social capital	2.97	0.90	.97	6
Financial well-being	3.03	0.83	.89	5
Psychological well-being	3.51	0.47	.87	11

Note. All scales use five-point.

Results

Perceived Sources of Job Insecurity (RQ1)

The findings point to a clear hierarchy in how this enclave's economic community experiences insecurity, and the top concerns are market-structural and competitive rather than individual or status-based. The three highest-rated sources, the broader economic environment ($M = 3.93$), competition from other workers ($M = 3.72$), and being undercut by lower-wage labor ($M = 3.67$), were each rated as a significant driver by roughly 70% of respondents. This clustering suggests that workers in these ethnic Chinese businesses locate the threat in the external labor market and in wage competition within their own job market, a pattern consistent with the enclave-economy literature's emphasis on thin margins, labor replaceability, and saturation in co-ethnic sectors (Sanders and Nee; Aldrich and Waldinger; Zhou, *Chinatown*).

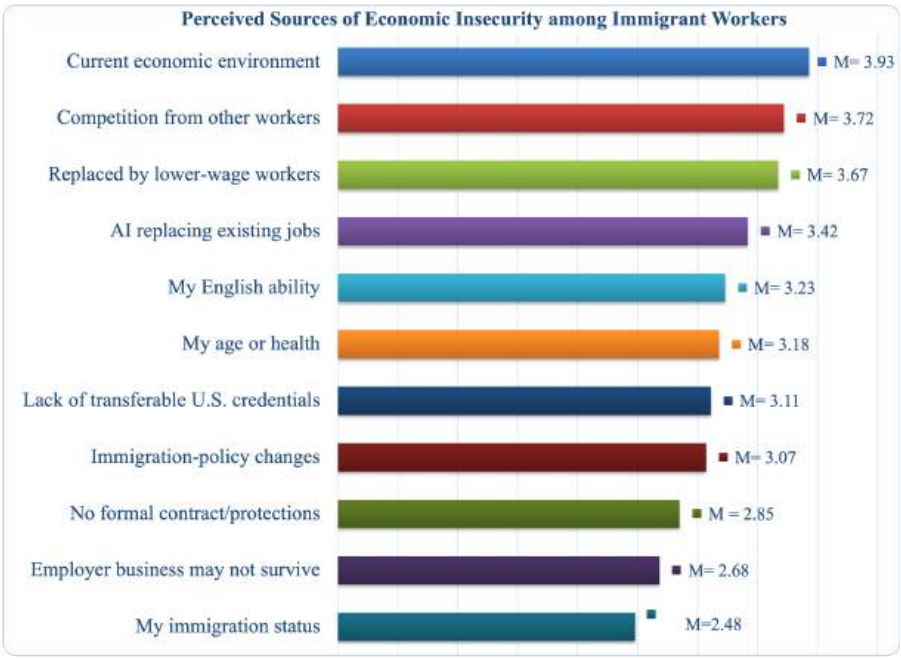
A second tier reflects human-capital and technological vulnerability: fear of AI displacement ($M = 3.42$), limited English ($M = 3.23$), age or health ($M = 3.18$), and lack of U.S.-recognized credentials ($M = 3.11$). The relatively high spread on English ability ($SD = 1.37$, the largest in the set) indicates that this is a polarizing concern, a real threat for some workers and a non-issue for others, which fits the bimodal English distribution in the sample (36.9% very good vs. 31.3% not very good or none).

Most notable is what sits at the bottom. Immigration status was rated the least influential of the eleven sources ($M = 2.48$), with 44 respondents reporting that it had no effect at all, and concern about immigration-policy changes ($M = 3.07$) fell below every human-capital and technological source. The two next-lowest sources, employer business viability ($M = 2.68$) and the absence of a formal labor contract ($M = 2.85$), are firm-level rather than status-based. Within this established community, in which 54.2% have lived in the United States for more than ten years, economic precarity appears to be experienced far more as a function of market and competitive forces than of legal or immigration vulnerability. Their insecurity, in other words, is

less about belonging in the country than about staying afloat in a competitive, low-margin labor market increasingly shadowed by automation. Figure 1 presents the full ranking.

Figure 1

Perceived Sources of Job and Economic Insecurity, Ranked (N = 179)



Note. Values are meant on a five-point scale (1 = *no effect at all*, 5 = *a very strong effect*).

Hypothesis Testing

Zero-order correlations among the study variables are presented in Table 2. Perceived job insecurity was negatively correlated with both financial well-being ($r = -.59$, $p < .001$) and psychological well-being ($r = -.55$, $p < .001$). Contrary to expectation, bonding social capital was weakly and negatively associated with financial well-being ($r = -.20$, $p < .01$) and was unrelated to psychological well-being ($r = -.13$, $p = .09$), whereas bridging social capital was positively associated with both outcomes ($r = .43$ and $r = .27$, respectively). Diagnostic checks indicated no violations of regression assumptions, and multicollinearity was not a concern (all variance inflation factors < 2.2).

Table 2

Zero-Order Correlations among Study Variables (N = 179)

Variable	1	2	3	4	5
1. Perceived job insecurity	-				
2. Bonding social capital	.12	-			
3. Bridging social capital	-.38***	-.17*	-		
4. Financial well-being	-.59***	-.20**	.43***	-	
5. Psychological well-being	-.55***	-.13	.27***	.53***	-

Note. * p < *.01.* * p* < .001.

Two hierarchical multiple regression models were estimated. In each, demographic control variables were entered in Step 1, and perceived job insecurity and bonding social capital were entered in Step 2.

Financial well-being. As hypothesized, perceived job insecurity was a significant negative predictor of financial well-being (β = -.39, *B* = −.44, *SE* = 0.07, *p* < .001), supporting H1. Bonding social capital, however, was not a significant predictor (β = -.10, *p* = .100), and H3 was therefore not supported. Among the covariates, older age (β = .26, *p* = .001) and higher English proficiency (β = .31, *p* < .001) were associated with greater financial well-being; education, years of U.S. residence, gender, and household income were not. Adding bridging social capital in Step 3 did not explain significant additional variance and bridging capital was not a significant predictor (β = .12, *p* = .167). Full results are included in Table 3.

Psychological well-being. Perceived job insecurity emerged as the strongest predictor and was negatively associated with psychological well-being (β = -.52, *B* = -.33, *SE* = 0.05, *p* < .001), supporting H2. Bonding social capital was not significantly associated with psychological well-being (β = -.08, *p* = .251), and thus H4 was not supported. Higher English proficiency was the only significant covariate (β = .23, *p* = .014). As with financial well-being, bridging social capital was not a significant predictor (β = .03, *p* = .799) either. Table 4 presents the full results.

Table 3

Hierarchical Regression Predicting Financial Well-Being (N = 179)

Predictor	**B**	**SE**	**β**	**p**
Step 1: Covariates				
Age	.18	.05	.26	.001
Gender (female)	-.16	.09	-.10	.088
Education	.01	.06	.01	.899
Years in the U.S.	.02	.07	.02	.783
English proficiency	.22	.09	.23	.019
Household income	.03	.04	.04	.508
Step 2:				
Perceived job insecurity	-.42	.07	-.38	< .001
Bonding social capital	-.16	.10	-.11	.091
Step 3:				
Bridging social capital	.11	.08	.12	.167

Table 4

Hierarchical Regression Predicting Psychological Well-Being (N = 179)

Predictor	**B**	**SE**	**β**	**p**
Step 1: Covariates				
Age	.02	.04	.04	.677

Gender (female)	-.08	.06	-.08	.216
Education	-.05	.04	-.11	.222
Years in the U.S.	-.06	.05	-.11	.215
English proficiency	.12	.06	.22	.059
Household income	-.02	.03	-.06	.386
Step 2:				
Perceived job insecurity	-.33	.05	-.52	< .001
Bonding social capital	-.07	.06	-.08	.250
Step 3:				
Bridging social capital	.01	.05	.03	.799

Discussion

In this study, I examined how immigrant workers in an ethnic enclave economy perceive the sources of their job insecurity, and how perceived job insecurity, co-ethnic bonding social capital, and cross-ethnic bridging capital relate to the financial and psychological dimensions of their well-being. Four findings have emerged from my analysis: the centrality of market-structural threats in workers' own accounts of their insecurity; the strong negative association between perceived job insecurity and both dimensions of well-being; the insignificant role of bonding social capital to predict either outcome; and the \ positive associations of bridging social capital and English proficiency.

Sources of Insecurity: Staying Afloat, Not Belonging

The findings from RQ1 locate the center of these workers' insecurity in the market rather than in the state. The three most influential sources, the broader economic environment, competition from other workers, and displacement by lower-wage labor, are all features of the competitive structure of enclave employment, consistent with the literature's characterization of co-ethnic niches as saturated, thin-margin sectors with high labor substitutability (Sanders and Nee; Aldrich and Waldinger; Zhou, *Chinatown*). A second tier reflects human capital and technological vulnerability, including displacement by artificial intelligence, limited English, age or health, and the absence of U.S.-recognized credentials, the last echoing Xie and Gough's distinction between enclave-oriented and mainstream-oriented human capital. Immigration status ranked lowest of the eleven sources, and concern about immigration-policy changes, though not the lowest-rated item, fell below every human-capital and technological source.

This ordering should be read considering who the respondents are. The sample consists mostly of "long-settlers", with the majority living in the United States for more than ten years, relatively educated, and drawn from an established enclave, and for such a community the pressing question is evidently not whether one may remain in the country but whether one can stay afloat in a crowded labor market increasingly shadowed by artificial intelligence and automation. The prominence of AI displacement in the middle tier is noteworthy: concerns usually documented among mainstream knowledge workers have reached the ethnic economy.

Job Insecurity as the Dominant Correlate of Well-Being

Hypotheses 1 and 2 were supported with strong statistical evidence. Perceived job insecurity was the single largest correlate of both financial well-being and psychological well-being, together with the covariates accounting for roughly half and one third of the variance in the respective outcomes. These associations are consistent with what existing literature suggest (Sverke, Hellgren, and Näswall; Cheng and Chan), which is also theoretically coherent: enclave employment represents informal employment relations, thin margins, and substitutable labor, under which the anticipation of job loss is chronic (Kwong, *The New Chinatown*; Sanders and Nee). The finding extends a robust occupational health literature to a population it has rarely examined-- first-generation immigrant workers surveyed in their own language, and indicates that the perceptual experience of insecurity, not only the objective wage penalty emphasized in the enclave debate, is a central dimension through which enclave employment influences workers' lives.

The Insignificance of Bonding Social Capital

Hypotheses 3 and 4 were not supported. Bonding social capital was unrelated to psychological well-being and, at the zero-order level, weakly negatively correlated with financial well-being. Given the theoretical expectation that dense co-ethnic ties supply both instrumental aid and belonging (Cohen and Wills; Keyes), this null finding may be explained by the follows.

The first is that the costs of co-ethnic solidarity offset its benefits. Portes and Sensenbrenner argue that the same closure that generates trust also imposes obligations to share resources, exerts pressure, and can serve as the very channel of within-group exploitation. On this account, a worker rich in co-ethnic ties is also a worker enmeshed in claims on her income and loyalty, and the small negative correlation with financial well-being is consistent with what the negative-social-capital literature suggests.

The second explanation is that there may be a distinction between getting by and getting ahead (Briggs; Putnam). Bonding ties circulate the resources a community already possesses; in a saturated ethnic labor market, what workers lack, namely access to better-paying employment, mainstream institutions, and non-redundant information, lies outside the group and is reached through bridging ties (Granovetter). The rest of the findings are are consistent with this account: bridging social capital was positively associated with both financial ($r = .43$) and psychological well-being ($r = .27$), a pattern closely reflecting Lancee's finding that immigrants' economic outcomes in the Netherlands returned to bridging but not bonding capital.

Acculturation Resources and the Mainstream Connection

English proficiency was the only covariate associated with both outcomes, and older age additionally predicted financial well-being. The English finding belongs to the same theoretical family as the bridging result: both proficiency and cross-group ties index connection to the mainstream opportunity structure, in Berry's terms an acculturation resource, and in Xie and Gough's terms mainstream-oriented rather than enclave-oriented capital.

Practical Implications

Findings from the study offer two practical implications. First, because perceived insecurity, not merely realized job loss, was the dominant correlate of distress, interventions that stabilize expectations, including clearer terms of employment, basic contractual protections, and accurate information about policy changes, may matter as much as wage levels. Second, community organizations serving enclave populations may serve members' well-being less by densifying within-group ties, which are already abundant, than by building bridges outward: language programs, credential-recognition assistance, and cross-community professional networks target exactly the resources this study found associated with better outcomes.

Limitations

Several limitations should be acknowledged in this study. First, the design is cross-sectional, so the associations presented here should not be interpreted as causation. For example, financially strained workers may perceive more insecurity in objectively identical jobs, and psychologically distressed workers may evaluate their prospects and networks more negatively. The associations reported here are consistent with, but cannot establish, the causal ordering the hypotheses imply. Second, the sample is a non-probability convenience sample of Chinese-speaking workers in one metropolitan area, recruited partly through personal networks, and is older, more settled, and more educated than many enclave workforces. Therefore, generalization to other communities, and especially to more recently arrived or more precarious ones, is not warranted. Third, all constructs were self-reported in a single instrument, so common method variance may inflate the observed associations (Podsakoff et al.). Fourth, the study observes only workers within the enclave; without a comparison group of co-ethnic workers in the mainstream economy, it can characterize variation among enclave workers but not the effect of enclave employment itself.

Future Directions

With the above limitations, here are a few ways that future research can continue. A comparison design sampling co-ethnic workers inside and outside the enclave would address the question this study could not, namely, whether enclave employment itself, and not only variation within it, shapes insecurity and well-being. Longitudinal measurement, such as through a panel survey, would allow the causal ordering of insecurity and well-being to be examined better. Finally, I mainly used a quantitative approach, but qualitative interviews could illuminate what the survey pattern can only gesture at: how workers themselves understand the obligations of co-ethnic ties, and whether the silence of bonding capital reflects its costs, its irrelevance to these outcomes, or its universality within the community.

Conclusion

Four decades of research have asked whether the ethnic enclave economy is a springboard or a trap, and have focused almost entirely on wages. This study collects primary

research data to understand workers from an enclave economy what makes their position feel insecure and how that insecurity, alongside their community ties, relates to how their lives are going. Their answers complicate both sides of the inherited debate. Their insecurity is real, market-made, and strongly tied to both their financial security and their psychological state. Meanwhile, the co-ethnic solidarity, which is at the center of the enclave concept, showed no significant associations with workers' financial or psychological well-being.

Works Cited

Aldrich, Howard E., and Roger Waldinger. "Ethnicity and Entrepreneurship." Annual Review of Sociology, vol. 16, 1990, pp. 111-35.

Berry, John W. "Immigration, Acculturation, and Adaptation." Applied Psychology: An International Review, vol. 46, no. 1, 1997, pp. 5-34.

Briggs, Xavier de Souza. "Brown Kids in White Suburbs: Housing Mobility and the Many Faces of Social Capital." Housing Policy Debate, vol. 9, no. 1, 1998, pp. 177-221.

Cheng, Grand H.-L., and Darius K.-S. Chan. "Who Suffers More from Job Insecurity? A Meta-Analytic Review." Applied Psychology: An International Review, vol. 57, no. 2, 2008, pp. 272-303.

Chin, Margaret M. Sewing Women: Immigrants and the New York City Garment Industry. Columbia UP, 2005.

Cohen, Sheldon, and Thomas A. Wills. "Stress, Social Support, and the Buffering Hypothesis." Psychological Bulletin, vol. 98, no. 2, 1985.

Consumer Financial Protection Bureau. Financial Well-Being: The Goal of Financial Education. Consumer Financial Protection Bureau, 2015.

De Witte, Hans. "Job Insecurity and Psychological Well-Being: Review of the Literature and Exploration of Some Unresolved Issues." European Journal of Work and Organizational Psychology, vol. 8, no. 2, 1999..

Diener, Ed, et al. "The Satisfaction with Life Scale." Journal of Personality Assessment, vol. 49, no. 1, 1985, pp. 71-75, https://doi.org/10.1207/s15327752jpa4901_13.

Goldberg, David P., and Paul Williams. A User's Guide to the General Health Questionnaire. NFER-Nelson, 1988.

Granovetter, Mark S. "The Strength of Weak Ties." American Journal of Sociology, vol. 78, no. 6, 1973, pp. 1360-80.

Hondagneu-Sotelo, Pierrette. Doméstica: Immigrant Workers Cleaning and Caring in the Shadows of Affluence. U of California P, 2001.

Keyes, Corey Lee M. "Social Well-Being." Social Psychology Quarterly, vol. 61, no. 2, 1998, pp. 121-40.

Kwong, Peter. Forbidden Workers: Illegal Chinese Immigrants and American Labor. New Press, 1997.

Lancee, Bram. "The Economic Returns of Immigrants' Bonding and Bridging Social Capital: The Case of the Netherlands." International Migration Review, vol. 44, no. 1, 2010, pp. 202-26.

Light, Ivan, and Steven J. Gold. Ethnic Economies. Academic Press, 2000.

Netemeyer, Richard G., et al. "How Am I Doing? Perceived Financial Well-Being, Its Potential Antecedents, and Its Relation to Overall Well-Being." Journal of Consumer Research, vol. 45, no. 1, 2018, pp. 68-89.

Pew Research Center. “Chinese Americans: Data on Asian Americans.” Pew Research Center, 2025, www.pewresearch.org/race-and-ethnicity/fact-sheet/asian-americans-chinese-in-the-u-s/. Accessed 31 July 2026.

Podsakoff, Philip M., et al. “Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies.” Journal of Applied Psychology, vol. 88, no. 5, 2003, pp. 879-903.

Portes, Alejandro. “Social Capital: Its Origins and Applications in Modern Sociology.” Annual Review of Sociology, vol. 24, 1998, pp. 1-24.

Portes, Alejandro, and Robert L. Bach. Latin Journey: Cuban and Mexican Immigrants in the United States. U of California P, 1985.

Portes, Alejandro, and Leif Jensen. “The Enclave and the Entrants: Patterns of Ethnic Enterprise in Miami before and after Mariel.” American Sociological Review, vol. 54, no. 6, 1989, pp. 929-49.

“What’s an Ethnic Enclave? The Case for Conceptual Clarity.” American Sociological Review, vol. 52, no. 6, 1987, pp. 768-71.

Portes, Alejandro, and Julia Sensenbrenner. “Embeddedness and Immigration: Notes on the Social Determinants of Economic Action.” American Journal of Sociology, vol. 98, no. 6, 1993, pp. 1320-50.

Portes, Alejandro, and Steven Shafer. “Revisiting the Enclave Hypothesis: Miami Twenty-Five Years Later.” The Sociology of Entrepreneurship, edited by Martin Ruef and Michael Lounsbury, Emerald, 2007, pp. 157-90. Research in the Sociology of Organizations, vol. 25.

Putnam, Robert D. Bowling Alone: The Collapse and Revival of American Community. Simon and Schuster, 2000.

Sanders, Jimmy M., and Victor Nee. “Limits of Ethnic Solidarity in the Enclave Economy.” American Sociological Review, vol. 52, no. 6, 1987, pp. 745-73.

Sverke, Magnus, Johnny Hellgren, and Katharina Näswall. “No Security: A Meta-Analysis and Review of Job Insecurity and Its Consequences.” Journal of Occupational Health Psychology, vol. 7, no. 3, 2002, pp. 242-64.

Vander Elst, Tinne, Hans De Witte, and Nele De Cuyper. “The Job Insecurity Scale: A Psychometric Evaluation across Five European Countries.” European Journal of Work and Organizational Psychology, vol. 23, no. 3, 2014, pp. 364-80.

Wilson, Kenneth L., and Alejandro Portes. “Immigrant Enclaves: An Analysis of the Labor Market Experiences of Cubans in Miami.” American Journal of Sociology, vol. 86, no. 2, 1980, pp. 295-319.

Xie, Yu, and Margaret Gough. “Ethnic Enclaves and the Earnings of Immigrants.” Demography, vol. 48, no. 4, 2011, pp. 12931315.

Zhou, Min. Chinatown: The Socioeconomic Potential of an Urban Enclave. Temple UP, 1992.

Zhou, Min, and John R. Logan. “Returns on Human Capital in Ethnic Enclaves: New York City’s Chinatown.” American Sociological Review, vol. 54, no. 5, 1989, pp. 809-20.

The 'Last Mile' of Skilling: How Social Enterprises Can Bridge the Gap Between Vocational Training and Actual Employment for People with Disabilities By Ayaan Jain

Abstract

Fragmented linkages between vocational training institutions, the private-sector, and persons with disabilities (PwDs) remain a major barrier to translating vocational training into actual employment. This paper examines whether social enterprises in India can address this "last-mile" problem by linking skilling, placement, and retention of PwDs. Using a comparative case-study approach, it analyzes two disability-focused social enterprises: Nabet India and Incluzza. Persistent barriers on both the supply and demand sides of the labor market are identified, including training curriculums, limited stakeholder engagement, inaccessible workplaces, weak implementation of disability inclusion measures, and inadequate support after placement. Findings suggest that social enterprises can improve these linkages, reducing ecosystem fragmentation more effectively than programs with limited integration. However, their impact is constrained by dependence on corporate partnerships, limited scalability in low-resource contexts, and lack of evidence on long-term retention. Stronger public frameworks and systematic tracking of post-placement outcomes are necessary to assess and expand their contribution to inclusive employment.

Keywords social enterprises, persons with disabilities, inclusive workforces, vocational training, employment

Introduction

Labor markets where Persons with Disabilities (PwDs) can participate on an equivalent basis with others are of primary importance. Employment of PwDs is critical to their social inclusion and economic independence. Inclusive employment opportunities are associated with higher incomes, better quality of life, and greater autonomy for PwDs, which translate into more productive workforces, reduce demands on welfare services, and create greater social integration and cohesion (Eurofound, 26; Rohwerder, 2). However, such inclusion remains complex. Disability not only encompasses different physical and mental experiences, but through its intersectionality, produces a heterogenous workforce that has largely been excluded from full participation within India's mixed economy (Limaye et al., 38–48).

Barriers to full and equal participation result from PwDs' interactions with several psychological and environmental barriers (World Health Organization). In India, the lack of education and training opportunities, financial resources, workplace environments, and employer attitudes have been identified as key barriers within the sector (Ramachandra et al.). PwDs also report negative attitudes of people, discrimination, harassment at workplaces, and information and communication barriers. Lack of accessible infrastructure, including public buildings and transportation, are also major hindrances (Khatoon). Even when employed, career progression

remains limited, with PwDs being more likely to be underemployed and underpaid (Hogan et al.).

Numbers reveal the scale. In India, on average, for all types of disability, 73.6% are out of the labor force (Shenoy, 2), while only 23% of women with disabilities are employed compared to their counterparts (Kumar et al.). Existing studies highlight that the type of disability is a crucial determinant of employment opportunities (Boman et al.). Gender and disability combined to double exclude women with disabilities from employment opportunities (Moral-Cabrero & Serrato-Calero, 1–6), which further intersect with variables like caste and residence (Khatoon). The inverse is also true: exclusion of PwDs imposes economic burdens on individuals, families, and societies at large. According to Buckup, nations lose up to 7% of their GDP due to the exclusion of PwDs (1).

From the legislative standpoint, India's Rights of Persons with Disabilities Act, (2016), which replaced the Persons with Disabilities Act of 1995, mandates that establishments, including those in the private sector, cannot discriminate against PwDs. However, top companies and multinationals in the private sector report dismal rates of inclusion (Shenoy, 2). Within government jobs, despite reservation quotas, systemic gaps persist till date, while schemes and policies have had limited success (Sharma). As a result, gains achieved through vocational education and skill development often fail to translate into meaningful opportunities.

Fragmented linkages between vocational training institutions, private-sector employers, and PwDs remain the central challenge within this ecosystem (Shenoy, 9). This paper examines whether social enterprises can bridge the last-mile delivery problem by analyzing the barriers that prevent trained PwDs from securing and sustaining employment, as well as analyzing the employment models adopted by two social enterprises in India, Nabet and Incluzza, and their effectiveness in linking vocational training to employment.

Background

Several state-level schemes, programs, and institutions provide vocational training to integrate PwDs into the mainstream workforce. The National Action Plan for Skill Development (NAP-SDP) aims for a conversion rate of 70% or higher with its training partners (Ministry of Social Justice and Empowerment, 11). Attempts to build linkages were also made through the PM-DAKSH–DEPwD Portal, which offers an integrated hub catering to the needs of PwDs. A network of non-governmental organizations, private trusts, and national councils also operate within this ecosystem. However, vocational programs involving the broader unemployed population reveal large systemic gaps within the framework: the Pradhan Mantri Kaushal Vikas Yojana (PMKVY), India's flagship skill certification scheme, trained nearly 14 million candidates between 2015 and 2023, yet only around 18% secured jobs (Ministry of Skill Development and Entrepreneurship).

Existing literature plots these challenges across demand and supply axes (Shenoy, 9). Companies are a major stakeholder in ensuring PwDs get linked to labor markets. Minimal private sector engagement severely restricts employment opportunities. Deterrents range from

attitudinal barriers of employers, perceived high investments for workplace adjustments, and a limited ecosystem of support organizations that help employers implement accessible infrastructure programs in workplaces (Venkatesh et al., 3142–3148). On the supply side, difficulties encompass insufficient levels of vocational training, non-standardized curriculums that do not align with market needs, competencies of training personnels themselves, and poor familial support networks (Naraharisetti & Castro, 1063; Venkatesh et al., 3142–3148).

Although no universally accepted official definition of Social Enterprises (SEs) exists, they are commonly understood as organizations driven by a social mission, guided by community-oriented values, and committed to reinvesting surpluses to advance objectives (Darko & Smith, 1–3). Social enterprises can play a crucial role in developmental contexts by enhancing employability through direct market linkages and by challenging negative perceptions. Both are especially relevant to India, where disability services and attitudes severely lag behind (Asian Development Bank, 18–19; Babikian & Hamdani, 40–46).

As discussed above, the ecosystem is presently characterized by high fragmentation among training providers, employers, NGOs, support organizations, and PwDs (Shenoy, 9). Hence, success depends on the ability of social enterprises to bridge these gaps as intermediary actors, and how effectively skilling, placement, workplace adaptation, and long-term retention are integrated across various stakeholders.

Discussion

Social enterprises follow a market-driven model where outcomes depend on creating viable business models, achieving financial sustainability, scaling up operations, and promoting transitions to the open labor market for PwDs. Founded in 2011, the National Association for the Blind (Employment & Training), or Nabet, is a Gurugram-based social enterprise working towards the empowerment of people with disabilities and people with visual impairments. Nabet explicitly links vocational training with private markets by positioning itself as a social enterprise supplying industry-ready manpower (Priya).

The enterprise developed its model through direct industry consultation. Beginning in 2009, founders approached local companies to identify skill requirements specific to sectors, and design curriculums that aligned with employer needs directly (‘Nabet India’). This led to the launch of a pilot batch of 15 visually impaired candidates in 2011 in collaboration with Idea Cellular (Aditya Birla Group). Over time, Nabet established partnerships with firms such as Mahindra Group, Vodafone, and Jeevansathi, integrating corporate participation into ideation, training, and placement processes.

Nabet also operates through two employment pathways. Companies either outsource functions to Nabet, with employees working from Nabet-managed centres, or hire candidates directly within their own workplaces. This creates a flexible model that accommodates varying levels of employer readiness and PwDs’ functional needs.

Nabet reports having trained more than 1,500 students across its vocational programs (‘Nabet India’). Reported outcomes also extend to addressing attitudinal barriers that frequently

limit employment, leading to greater confidence, self-belief, and readiness for responsible corporate assignments among PwDs (SRV Media). Nabet is also attempting to address the information-access barriers and collaborate with the Ministry of Social Justice and Empowerment in doing so (Priya). Summarily, salient features include addressing ecosystem fragmentation, aligning curricula with industry demand, connecting vocational training with corporate recruitment pathways, and building confidence and workplace readiness among PwDs.

Established in 2011, Incluzza was created to extend EnAble India's existing work on disability inclusion and livelihoods into a social enterprise oriented model (Incluzza). In 2021, it launched Incluzza as a comprehensive, one-stop solution for disability and inclusion requirements, shifting the focus from training PwDs alone to facilitating direct linkages for PwDs. Incluzza addresses several demand side, or employer-side barriers that interrupt the transition to employment for PwDs. Its intervention combines physical, digital, and content accessibility support for companies, helping workplaces become RPwD Act compliant, and leadership training for managers responsible for recruitment and workplace inclusivity. In doing so, it addresses four major barriers to disability employment: inaccessible infrastructure, information and communication exclusion, weak implementation of legal obligations, and limited managerial capacity to accommodate and retain PwDs. While limited data is available on its scale, it has worked with major companies including Dell Technologies, Thomas Reuters, CISCO, among others, and its parent, EnAble India, has worked with more than 500,000 stakeholders indicating scope for substantial expansion (Zero Project).

Both Nabet and Incluzza illustrate that the integration of PwDs into labor markets is a dynamic process that involves interactions between organizational (demand side), personal (supply side), and institutional factors (Eurofound, 1). Results align with findings that market-based approaches tend to have a clearer effect on transitions to the open labor market. A key reason is that directly working with PwDs gives social enterprises a greater insight into the heterogeneity of lived experiences, knowledge, and skills, while also fitting these gains within existing systems and power structures (Mcintosh & Wright, 449–467). Hence, within a mainstreaming approach, and through wider uptake and reach, they contribute significantly to labor-market inclusion of vulnerable groups.

Nabet's strengths lie in its direct consultation with employers, which reduces the mismatch between training curriculums and industry requirements. Through two employment pathways, it is also able to accommodate employers with different levels of readiness. Incluzza operates at a different level. Rather than primarily preparing PwDs for work, it focuses on making organizations capable of employing PwDs. Its intervention combines accessibility support and leadership training for management and recruitment teams for better workplace inclusion. This heterogeneity reflects differing approaches, both of which are relevant insofar as they address barriers that government and vocational programmes often leave unresolved.

However, several limitations must also be noted. Inclusive integration does not end once employment is secured. Without effective long-term integration processes, people with disabilities often face mistreatment in workplaces and continue to be affected by prejudices held

against them (Shenoy, 8). Scalability also remains uncertain, as both cases depend heavily on corporate partnerships, which may be difficult to replicate in rural areas where demand, employer networks, and accessible infrastructure are far more limited. As Chui et al. (176–196) warn, this could result in "cocooned inclusion" rather than "actual transition". It can also reinforce normative demarcations between abled workers and PwDs, rather than reforming mainstream employment systems, and may inadvertently deflect attention from institutionalizing proactive disability policy measures. Moreover, Eurofound argues that exclusion of PwDs stems from macro-level factors shaped by economic systems, differing levels of economic development within a territory, and the availability of resources, all of which fall beyond the scope of what a social enterprise alone can address (14). Hence, while social enterprises can strengthen the last-mile between vocational training and employment, they should not be seen as substitutes for systemic reforms, but as complementary strengths within wider disability-inclusion policy and practices.

Conclusion

This paper examined the role of social enterprises in bridging the last mile gap between vocational training and meaningful employment for PwDs in India. Findings show that the principal weakness in the current ecosystem is the fragmented connection between training providers, employers, organizations, and PwDs. Large-scale vocational programs have demonstrated that training alone does not guarantee employment, highlighting scope for intervention by social enterprises.

Together, both cases suggest that social enterprises can improve alignment between skilling and recruitment, make workplaces more accessible, and address gaps that are often weakly connected within conventional vocational training systems. Research on the role of social enterprises in improving PwDs livelihoods is still limited globally and within India, particularly in relation to long-term outcomes. A lack of data on socioeconomic correlates of disability also leaves significant gaps in understanding results through an intersectional framework. Publicly available data on retention, wage progression, workplace accommodation, and career advancement would be critical in assessing long-term effectiveness and relevance of social enterprises within this landscape.

Future research can examine outcomes across different disability categories, women with disabilities, caste intersections, rural populations, and other forms of structural disadvantage. On a policy level, public awareness of the role of social enterprises, clear implementation guidelines and operational frameworks, and systematic tracking of long-term outcomes would improve accountability within the ecosystem. Overall, social enterprises can play a significant intermediary role in connecting vocational training with meaningful employment, but they cannot be seen as substitutes for broader reforms in labor market inclusion.

Works Cited

Asian Development Bank. India Social Enterprise Landscape Report. 2nd ed., 2012.

Babikian, V. Armineh, and Yani Hamdani. "Social Enterprises and Transition to Employment for People Labeled with Intellectual and Developmental Disabilities". Current Developmental Disorders Reports, vol. 10, no. 1, 2023, pp. 40–46. https://doi.org/10.1007/s40474-023-00267-7.

Boman, Tomas, et al. "Employment Opportunities for Persons with Different Types of Disability". Alter, vol. 9, no. 2, 2015, pp. 116–29. https://doi.org/10.1016/j.alter.2014.11.003.

Buckup, Sebastian. The Price of Exclusion: The Economic Consequences of Excluding People with Disabilities from the World of Work. International Labour Office, Skills and Employability Department, 2009. Employment Working Paper, no. 43.

Chui, Cheryl Hiu-Kwan, et al. "The Role of Social Enterprises in Facilitating Labour Market Integration for People with Disabilities: A Convenient Deflection from Policy Mainstreaming?" Journal of Social Policy, vol. 52, no. 1, 2023, pp. 176–96. https://doi.org/10.1017/S0047279421000490.

Darko, Emily, and William Smith. Social Enterprise: Constraints and Opportunities - Evidence from Vietnam and Kenya. Overseas Development Institute, 2014.

Eurofound. Disability and Labour Market Integration: Policy Trends and Support in EU Member States. Publications Office of the European Union, 2021. https://data.europa.eu/doi/10.2806/49656.

Hogan, Anthony, et al. "Workforce Participation Barriers for People with Disability". International Journal of Disability Management, vol. 7, 2012, pp. 1–9. https://doi.org/10.1017/idm.2012.1.

Incluzza. Incluzza – By Enable India Solutions. https://incluzza.org/. Accessed 31 July 2026.

Khatoon, Farhana. "Employment Status of Persons with Disabilities in India: Insights from National Sample Survey Data 2018". Social Sciences & Humanities Open, vol. 13, 2026. https://doi.org/10.1016/j.ssaho.2026.102804.

Kumar, Sunaina, et al. Leaving No One Behind: Strategies for Disability Inclusion in India. ORF Special Report no. 240, Observer Research Foundation, 2024, https://www.orfonline.org/research/leaving-no-one-behind-strategies-for-disability-inclusion-in-india.

Limaye, Sandhya, et al., editors. Disability as Diversity in India: Theory, Practice, and Lived Experience. Routledge, 2024.

Mcintosh, Ian, and Sharon Wright. "Exploring What the Notion of 'Lived Experience' Offers for Social Policy Analysis". Journal of Social Policy, vol. 48, no. 03, 2019, pp. 449–67. https://doi.org/10.1017/S0047279418000570.

Ministry of Skill Development and Entrepreneurship. "State-Wise and District- Wise Candidates Trained and Placed under PMKVY from 2015 to 31.03.2023". 2023.

Ministry of Social Justice and Empowerment. “National Action Plan for Skill Development of Persons with Disabilities”. 2023, https://cdnbbsr.s3waas.gov.in/s3e58aea67b01fa747687f038dfde066f6/uploads/2023/12/20240509194998763.pdf.

Moral-Cabrero, Eva, and Ma Mercedes Serrato-Calero. “Feminist Disability Studies”. The Palgrave Encyclopedia of Disability, edited by Gabriel Bennett and Emma Goodall, Springer Nature Switzerland, 2025, pp. 1–6. https://doi.org/10.1007/978-3-031-40858-8_472-1.

‘Nabet India’. Nabet India, https://nabetindia.org/. Accessed 31 July 2026.

Naraharisetti, Ramya, and Marcia C. Castro. “Factors Associated with Persons with Disability Employment in India: A Cross-Sectional Study”. BMC Public Health, vol. 16, no. 1, 2016, p. 1063. https://doi.org/10.1186/s12889-016-3713-6.

Priya, Lekshmi. “This Organisation Helps Train Differently Abled People to Find Employment”. The Better India, 2017. https://thebetterindia.com/108090/training-and-employment-for-diffrently-abled-nabet-manesar/.

Ramachandra, Srikrishna S., et al. “Factors Influencing Employment and Employability for Persons with Disability: Insights from a City in South India”. Indian Journal of Occupational and Environmental Medicine, vol. 21, no. 1, 2017, p. 36. https://doi.org/10.4103/ijoem.IJOEM_44_16.

Rohwerder, Brigitte. “Disability Inclusion: Topic Guide”. GSDRC, University of Birmingham., 2015. https://gsdrc.org/topic-guides/disability-inclusion/.

Sharma, Nidhi. “Parliamentary Panel Seeks Compliance Data on Disability Quota in Government Jobs”. The Economic Times, 22 July 2026. The Economic Times - The Times of India, https://economictimes.indiatimes.com/news/india/parliamentary-panel-seeks-compliance-data-on-disability-quota-in-government-jobs/articleshow/132545959.cms?from=mdr.

Shenoy, Meera. “Persons With Disability & The India Labour Market: Challenges And Opportunities”. International Labour Organization, 2011, https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@asia/@ro-bangkok/@sro-new_delhi/documents/publication/wcms_229259.pdf.

SRV Media. Nabet India’s Initiative to Empower Persons with Disability with IT/ITES Skills. 2022.

The Rights of Persons with Disabilities Act. Act No. 49, 2016.

Venkatesh, Kundurthi, et al. “Vocational Skill Training Programs for Persons with Intellectual Disability (PID) and Trainers’ Perspective during and Post Vocational Skill Training”. Journal of Family Medicine and Primary Care, vol. 12, no. 12, 2023, pp. 3142–48. https://doi.org/10.4103/jfmpc.jfmpc_433_23.

World Health Organization. World Report on Disability 2011. World Health Organization, 2011.

Zero Project. Enable India | Zero Project. https://zeroproject.org/view. Accessed 31 July 2026.

Banking System.” Economics Letters, vol. 250, 2025, article 112303.

Mapping Access to Urban Green Infrastructure in Bengaluru By Sanya Pai

Abstract

In my project, I analyse whether green infrastructure is equally available across selected Bengaluru neighborhoods. The study uses OpenStreetMap data through OSMnx to compare parks, lakes, schools, hospitals, clinics, bus stops, and metro or rail stations around eight neighborhoods: Indiranagar, Jayanagar, Koramangala, Malleshwaram, Electronic City, Whitefield, Hebbal, and Yelahanka. A Green Access Score is created using park availability, distance to the nearest park, lake proximity, and supporting amenity access. The results show that green access is not equal. Indiranagar ranks highest, followed by Jayanagar and Koramangala, while Hebbal and Yelahanka score lower in the center-point method. The project is exploratory, but it shows how public data can help students study a real planning problem.

1. Introduction

This project studies urban green infrastructure in Bengaluru. Green infrastructure means parks, lakes, gardens, playgrounds, urban forests, and other open spaces that make city life healthier and more livable. These places matter because they give people space to walk, play, rest, and cool down. The World Health Organization explains that urban green spaces can support mental health, physical activity, social connection, and protection from heat, noise, and air pollution.

Bengaluru is a useful city for this question because it is growing quickly and still has a strong identity as a city of parks and lakes. At the same time, many of its lakes and green spaces are under pressure from rapid urbanisation. AP News has reported that Bengaluru once had more than 250 lakes in the 1970s, while only about 180 remain today, with many in poor condition. This makes the city a good case for asking whether green access is spread fairly across different neighborhoods.

The aim is not to make an official government map. The aim is to build a simple and transparent student project that can be understood and repeated. The main idea is to use the same type of public map data for each neighborhood and then compare which areas appear to have stronger access to parks and lakes.

2. Research Question

The main research question is: how evenly is green infrastructure distributed across selected Bengaluru neighborhoods, and what planning patterns can be seen from that distribution?

The project also asks which selected neighborhoods have the best and worst mapped access to parks and lakes, whether older or more planned neighborhoods appear to have better local green access, and what planners should keep in mind when using public map data to understand green access.

3. Study Area

The study uses eight Bengaluru neighborhoods. Indiranagar and Koramangala are treated as affluent mixed-use areas. Jayanagar and Malleshwaram represent older dense residential areas. Electronic City and Whitefield represent technology corridors. Hebbal is a mixed-use northern area, while Yelahanka represents a more peripheral northern area.

These places were selected because they represent different types of urban growth. Some are older and more planned, some are commercial and residential, some are technology corridors, and some are northern or peripheral growth areas. The selection is not meant to represent all of Bengaluru. It is meant to give enough variety for a useful comparison.

Each neighborhood is represented by a center point. This makes the project easier to repeat, but it also means the results depend on the chosen point. For large areas such as Yelahanka, Hebbal, Whitefield, and Electronic City, a different point could change the measured access.

4. Data

The main data source is OpenStreetMap, maintained by OpenStreetMap contributors. The project uses the Python package OSMnx, which Boeing describes as a tool for working with graph-theoretic OpenStreetMap street networks (215). For each neighborhood, the pipeline starts with a fixed latitude-longitude center point from data/neighborhoods_seed.csv. The script then queries OpenStreetMap features within a 2 km radius.

The extracted feature categories include parks, playgrounds, lakes and reservoirs, urban forests, schools, hospitals, clinics, bus stops, metro or rail stations, and public transport points. These categories help connect environmental access with everyday urban services.

OpenStreetMap is useful because it is public and reproducible. The same query can be checked and rerun. At the same time, OpenStreetMap is created by contributors, so coverage may be better in some places than in others. A Nature Communications paper by Herfort and others also notes that OpenStreetMap is widely used for urban analysis, but analysts must be careful about uneven spatial coverage.

The reproducible script is scripts/collect_osm_data.py. The main processed dataset is data/processed/neighborhood_summary.csv. The project also generates an interactive map at outputs/maps/bengaluru_green_access_map.html.

5. Data Cleaning

The first automated data pull gave useful results, but the data had to be cleaned before being used in the report. Some OpenStreetMap features appeared in more than one neighborhood query because nearby neighborhoods overlap. The script now deduplicates features by OSM element, id, and category.

Broad water tags also included drains and small unnamed water fragments. The method now counts lakes only when the feature appears to be a lake or reservoir, or when the name

indicates a lake or kere. Metro station entrances were also initially counted separately from stations, so the method now counts metro or rail stations rather than every station entrance.

These cleaning choices make the report more believable. If every drain were counted as a lake, or every metro entrance were counted as a station, the final score would be misleading.

6. Methodology

For each neighborhood, the analysis calculates park count within 1 km, playground count within 1 km, lake count within 2 km, distance to the nearest park or urban forest, distance to the nearest lake, school count within 1 km, hospital count within 1 km, clinic count within 1 km, bus-stop count within 1 km, metro or rail station count within 2 km, and distance to the nearest metro or rail station.

The main index is the Green Access Score. It ranges from 0 to 100 and combines four parts:

Green Access Score =
35% park availability score
+ 30% distance-to-nearest-park score
+ 25% lake/water-body proximity score
+ 10% public amenity balance score

Higher park counts increase the score. Shorter distances to parks and lakes increase the score. Amenities get a smaller weight because the project is mainly about green access, not general urban infrastructure.

The index is kept simple on purpose. A more advanced project could use official ward boundaries, population density, walking-network distance, park size, tree cover, and field surveys. For this school-level project, a transparent and explainable method is more suitable.

7. Results

The cleaned ranking is shown below.

Rank	Neighborhood	Green Access Score
1	Indiranagar	72.22
2	Jayanagar	58.28
3	Koramangala	57.68
4	Malleshwaram	47.61
5	Electronic City	37.77
6	Whitefield	36.62
7	Hebbal	17.61
8	Yelahanka	13.09

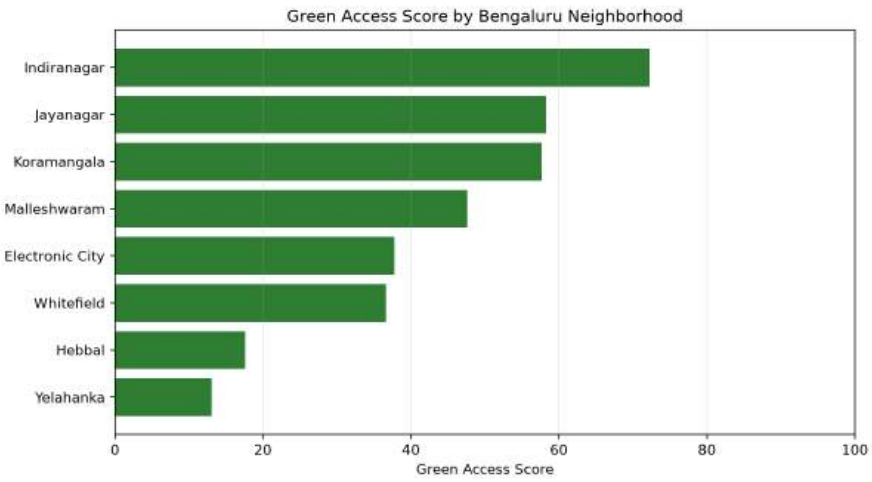


Figure 1. Green Access Score by Bengaluru Neighborhood

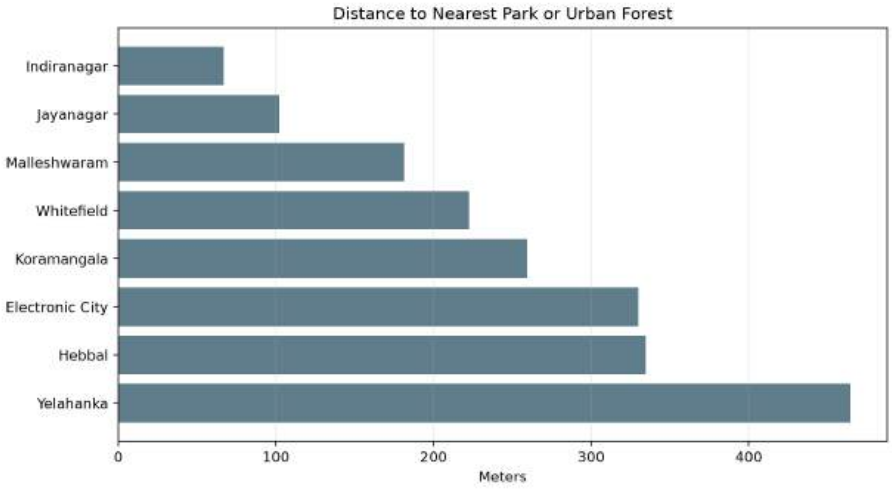


Figure 2. Distance to Nearest Park or Urban Forest

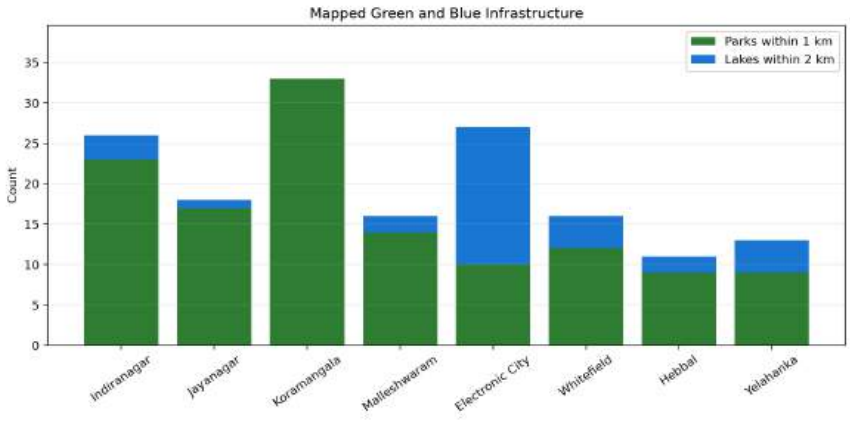


Figure 3. Mapped Green and Blue Infrastructure Counts

The results show that Indiranagar has the strongest combined access profile among the selected neighborhoods. It has many mapped parks within 1 km, a short distance to the nearest mapped park, and strong amenity access. Jayanagar also performs well because of short park distance and balanced amenities.

Koramangala has the highest park count within 1 km, but its score falls because lake proximity from the selected center point is weaker. This shows that green access is not only about counting parks. A neighborhood can have many parks and still score lower on blue-infrastructure access.

Electronic City has many mapped lakes within 2 km, but weaker park and amenity values reduce its overall score. Whitefield has a similar middle position. Hebbal and Yelahanka score lower in this method. This does not mean they have no lakes or parks. The map captures Hebbal Lake, Nagawara Lake, Yelahanka Lake, and Jakkur Lake. The lower score mainly means that, around the selected center points, park density and access balance are weaker.

8. Old / Planned Bengaluru and Newer Growth Comparison

The project also compares two broad groups. The old/planned Bengaluru sample includes Indiranagar, Jayanagar, Koramangala, and Malleshwaram. The newer-growth sample includes Whitefield, Electronic City, Hebbal, and Yelahanka.

This grouping is not an official historical boundary. Hebbal and Yelahanka have older settlement histories, but they are included in the newer-growth sample because the selected points represent Bengaluru's northern expansion and transport-linked growth. The purpose is only to ask whether the selected older/planned sample appears to have stronger local green access than the selected newer-growth sample.

Group	Avg. Green Access Score	Avg. Parks within 1 km	Avg. Lakes within 2 km	Avg. Nearest Park (m)
Old / planned Bengaluru sample	58.95	21.75	1.50	152.4
Newer growth Bengaluru sample	26.27	10.00	6.75	337.9

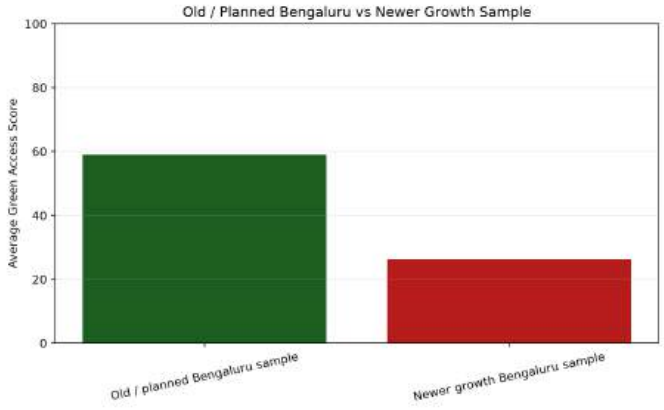


Figure 4. Average Green Access Score: old/planned Bengaluru sample vs newer growth sample

The comparison supports this idea. The old/planned Bengaluru sample has an average Green Access Score of about 58.95, while the newer-growth sample has an average score of about 26.27. The old/planned group also has a higher average number of parks within 1 km. This suggests that the older sampled neighborhoods have a denser local distribution of smaller green spaces.

The newer-growth group has a higher average lake count within 2 km, mainly because Electronic City, Yelahanka, Whitefield, and Hebbal are closer to larger lake systems. But this does not automatically mean stronger daily access, because the index also gives importance to walkable park access and distance to the nearest park.

9. Spot Check

A spot check was done to see whether the OpenStreetMap pipeline was capturing recognizable places. In Indiranagar, the mapped features include Halasuru Lake, Defence Colony Park, Indiranagar Park, Kuvempu Park, Jogger's Park, and Indiranagar metro station. In Koramangala, the pipeline captures ST Bed Park, Wipro Park, Lakshmi Devi Park, Udupi Garden Park, and Agara Lake. In Yelahanka, it captures Yelahanka Lake Park, Allalasandra Lake Park, Jakkur Lake Park, Yelahanka Lake, and Puttenahalli Lake. In Hebbal, it captures Hebbal Lake, Nagawara Lake, and Lumbini Gardens.

This makes the automated method more credible because the results are not just abstract categories. However, the spot check also shows one caution: public map data can show that a park or lake exists, but it cannot show whether the place is clean, safe, open, well maintained, or easy to enter.

10. Discussion and Policy Recommendations

The main finding is that mapped green access varies a lot across the selected neighborhoods. Older and centrally located neighborhoods such as Indiranagar and Jayanagar perform well in the index, while larger or newer-growth areas are more sensitive to the exact center point chosen.

One explanation is that small, distributed parks matter. Indiranagar and Jayanagar do well not only because they are near major ecological features, but because they have many smaller parks close to everyday residential life. This supports the idea that planning should not only focus on famous parks or big lakes. It should also focus on the smaller open spaces people can actually reach.

Another finding is that large lakes do not automatically create equal green access. Hebbal and Yelahanka have important lakes nearby, but that does not always translate into evenly distributed, walkable green access around the selected points. Research on Bengaluru lakes also supports this concern. Murphy, Enqvist, and Tengö show that lake stewardship in Bangalore involves both ecology and social access, while Mandal and Manasi describe how civic groups have had to work with public agencies to save and manage Bengaluru lakes.

The newspaper sources support the same broader issue. The Indian Express report on Yelahanka Lake shows how restoration can create a jogging track and children's park that people actually use. Deccan Herald's article on neighbourhood lakes argues for planning the area around lakes, not just the water body itself. Times of India has reported large-scale lake encroachments in Bengaluru Urban, which shows why lake protection is also a land-use issue.

The first policy recommendation is that fast-growing and peripheral areas should not rely only on large lakes or isolated parks. They need smaller, distributed green spaces that are walkable from residential areas. The second recommendation is that lake restoration should include access planning. A restored lake is more useful when people can reach it safely through footpaths, crossings, public transport, shaded streets, and clear entrances.

The third recommendation is that planning authorities should monitor green access at the neighborhood scale. City-level totals can hide local inequality. A city may have many parks overall while some neighborhoods remain poorly served within walking distance. The fourth recommendation is that technology corridors such as Whitefield and Electronic City should treat green access as basic infrastructure. Roads, drainage, housing, and offices are not enough by themselves. High-growth areas also need ecological and recreational infrastructure that residents can use in daily life.

The cautions is also important. OpenStreetMap completeness varies by area. The neighborhoods are represented by center points and circular radii, not official boundaries. The score measures mapped proximity and counts, not park size, maintenance, safety, biodiversity, or actual usage. Therefore, the Green Access Score should not be treated as an official planning metric. It is a simple comparison tool that helps reveal patterns.

11. Conclusion

In my project, public geospatial data is used to study urban green infrastructure in a simple and reproducible way. Among the selected Bengaluru neighborhoods, green access is uneven. Indiranagar, Jayanagar, and Koramangala score highest, while Hebbal and Yelahanka score lower in the center-point radius method.

The biggest lesson is that green access depends on both presence and distribution. A city may have lakes and parks, but what matters for residents is whether those spaces are close, usable, safe, and connected to everyday life. For Bengaluru, this means that protecting large lakes is important, but so is building and maintaining smaller neighborhood-level green spaces.

Works Cited

Boeing, Geoff. ‘OSMnx: A Python Package to Work with Graph-Theoretic OpenStreetMap Street Networks’. The Journal of Open Source Software, vol. 2, no. 12, Apr. 2017, p. 215, https://doi.org/10.21105/joss.00215.

Herfort, Benjamin, et al. ‘A Spatio-Temporal Analysis Investigating Completeness and Inequalities of Global Urban Building Data in OpenStreetMap’. Nature Communications, vol. 14, no. 1, July 2023, https://doi.org/10.1038/s41467-023-39698-6.

Kumar Pateel, Pavan. ‘Encroachments Gobble up 730 of 837 Lakes in Bengaluru Urban: Survey’. The Times of India, 11 Jan. 2026, https://timesofindia.indiatimes.com/city/bengaluru/encroachments-gobble-up-730-of-837-lakes-in-bengaluru-urban-survey/articleshow/126456722.cms.

Mandal, Dipak, and Manasi S. ‘How Civic Groups Are Meeting the Challenges of Saving Bengaluru Lakes: A Study’. Institute for Social and Economic Change Working Paper, vol. 528, Dec. 2021, https://www.isec.ac.in/wp-content/uploads/2023/07/WP-528-Dipak-Manadal-and-S-Manasi-Final-1.pdf.

Murphy, Ailbhe, et al. ‘Place-Making to Transform Urban Social–Ecological Systems: Insights from the Stewardship of Urban Lakes in Bangalore, India’. Sustainability Science, vol. 14, no. 3, Feb. 2019, pp. 607–23, https://doi.org/10.1007/s11625-019-00664-1.

OpenStreetMap. ‘OpenStreetMap’. OpenStreetMap, 2024, https://www.openstreetmap.org/.

Rahi, Aijaz, and Sibi Arasu. ‘As Climate Risks Grow, India’s Bengaluru Is Trying to Save Its Vanishing Lakes’. AP News, Associated Press, 20 Sept. 2025, https://apnews.com/article/climate-change-lakes-bengaluru-india-water-restoration-3d5f548dda4176a9465efb3940104214.

Suresh, Vardhini, et al. ‘Reimagining the Ring around the Neighbourhood Lakes’. Deccan Herald, 28 Mar. 2024, https://www.deccanherald.com/india/karnataka/bengaluru/reimagining-the-ring-around-the-neigbourhood-lakes-2955245.

Thakur, Aksheev. ‘Lakes of Bengaluru: Locals Throng Yelahanka Lake after Civic Body Breathes Life into It’. The Indian Express, 19 Mar. 2023, https://indianexpress.com/article/cities/bangalore/lakes-of-bengaluru-locals-throng-yelahanka-lake-civic-body-breathes-life-8505241/.

World Health Organization. Urban Green Spaces and Health. World Health Organization Regional Office for Europe, 1 Oct. 2016, https://www.who.int/europe/publications/i/item/WHO-EURO-2016-3352-43111-60341.

Exploring the Role of Dual N-Back Training in an Academic Setting Among the Pre-Adolescent Population By Tamanna Sethia

Abstract

Working memory is a core component of executive functioning essential for learning, reasoning, holding, and manipulating information. Previous research has focused on finding how to use tasks like the Dual N-Back to train and improve working memory and to study its effects as near or far transfer of skills. In this study, we explore the effects of Dual N-Back training on the adolescent population aged 9 through 13 in a school-based setting. We use a mixed-method approach to assess both working memory capacity progression and behavioural transfer effects. Quantitative evidence shows a general trend of task improvement, though a significant plateau at the 3-back level is observed. Qualitative evidence points out self-reported benefits across four themes: enhanced attentional control and distractor inhibition, subjective expansion of working memory, increased academic self-efficacy, and heightened awareness of environmental modulators. Our findings suggest that adaptive working memory training using tasks like Dual N- Back results in benefits beyond just maximizing working memory capabilities; it improves metacognitive awareness and promotes individual learning strategies to keep up with the repetitive, cognitively demanding tasks. This, in turn, provides psychological scaffolding and academic confidence to the student. Further studies should be conducted to evaluate a way to integrate such training in the educational curriculum that improves foundational executive functioning and promotes long-term cognitive health.

Keywords

Working memory, Dual N-Back, Brain Training, Cognitive Training, Attentional Control, Metacognition, School-aged Children, Executive Control

1.0 Introduction

Our memory is essential to our existence. It is fundamental to human cognition, allowing us not only to remember the what, how, and when of daily activities but also to actively hold and manipulate relevant information through working memory. Working memory is a core component of executive functioning. It supports manipulation and temporary maintenance of information and is crucial for learning, reasoning, decision-making, as well as goal-directed behavior (Baddeley & Hitch, 1974; Baddeley, 1992).

Growing up in our current digital age involves constant exposure to digital devices and multiple streams of highly engaging content with intense attentional demands. These streams of digital multimedia environments often encourage rapid switching between different sources. Studies have revealed how executive functioning, working memory capacity, and attentional control are affected due to excessive media use, resulting in lower academic performance (Cain et al., 2016). Others have also stated how media multitasking may affect components of working memory, like sustained attention and information updating (Uncapher et al., 2017). These

observations may suggest that the current educational paradigms may not be optimal in supporting the development of foundational cognitive capacities like executive functioning for effective learning. Researchers have found that working memory capacity continues to mature throughout childhood and adolescent developmental stages and is accompanied by executive function development. This makes this developmental stage particularly important for interventions aimed at strengthening cognitive performance (Shinomiya et al., 2015).

Training working memory has been explored extensively as a method of improving foundational cognitive abilities. Researchers have increasingly explored computerized training tasks designed to strengthen working memory. One of the most popular paradigms is the Dual N-Back task, which is an adaptive memory training program that simultaneously engages the visuospatial and auditory working memory (Jaeggi et al., 2008). The task involves two independent sources of information: one is visual, and the other is auditory. The participants have to identify whether either stimulus matches the corresponding stimulus presented N trials earlier. As N increases in number, so does the difficulty; its adaptable nature makes it suitable as a training paradigm. The effectiveness of Dual N-Back training has been studied extensively post Jaeggi et al. (2008). The current literature supports that repeated practice leads to substantial improvements in the memory task itself, but the evidence of this improvement on other broad factors like academics or intelligence is mixed (Melby-Lervåg & Hulme, 2013; Soveri et al., 2017).)

In this current paper, we explore the effects of Dual N-back training in an educational setting, although executive functioning and its role in academia have been studied previously; school curricula largely overlook systematic working memory training. Our study will address this gap by studying school-based cognitive interventions and checking their role in possibly strengthening attention and working memory, which may potentially support educational and long-term cognitive health.

2.0 Methods

2.1 Research Design

Our study employed a mixed-method longitudinal design to evaluate the effectiveness of adaptive Dual N-Back working memory training among school-aged children. We collected quantitative data on each session, as well as a qualitative survey, before the initiation of the training program and after its final session. The quantitative component consisted of training performance data collected over 15-16 consecutive days, with one session per day and 20 adaptive training blocks. The qualitative component consisted of pre and post-training surveys that explored participants' motivation, learning experience, perceived cognitive changes, and perception of the training itself.

2.2 Participants

Our study included 42 students aged 9 to 13 enrolled in grades 5 to 8. Participants were selected via convenience sampling. Across all participants, the surviving dataset of Dual N-Back training has 640 completed training sessions and 13,100 individual records. This dataset provided us with a comprehensive longitudinal insight into the effects of Dual N-Back.

2.3 Material

2.3.1 Dual N-Back Task

We used an adaptive Dual N-Back task for daily working memory training. This is a cross-modal working memory paradigm where participants are exposed to two independent streams of information simultaneously: Visual stimuli (a blue square moving across a 3x3 grid) and auditory stimuli (spoken letters played via audio).

The Objective: A "match" occurs when the current position of the square, the spoken letter, or both, is identical to the stimuli presented exactly N steps backward in the sequence.

N = 1 (Baseline Attentional Control): The participant compares the current visual/auditory stimulus to the one immediately preceding it (1 step back).

N = 2 (Working Memory Capacity - Standard Challenge): The participant must hold the current stimuli in mind while comparing them to what appeared 2 steps ago

N = 3+ (High Cognitive Load & Executive Function): The participant tracks matches 3 steps backward.

The levels scaled from N=1 and were adaptively scaled with respect to the abilities of the student participating; the maximum level reached was N=8.

2.3.2 Surveys

Participants were administered a questionnaire before commencing the training session on day 1 and after completion of the training on day 15/16. The initial survey collected demographic information, prior exposure to cognitive training, initial motivation, and participants' understanding of the programme. The post-intervention survey assessed perceived cognitive changes, motivation, learning experience, environmental facilitators and barriers, and overall perceptions of the intervention.

2.4 Data Analysis

The longitudinal data from Dual N-back training were used to perform the quantitative analysis. We analysed the quantitative data using Python 3.

We analysed the qualitative responses from the pre- and post-intervention surveys using reflexive thematic analysis following the six-phase framework proposed by Braun and Clarke (2006). We familiarized ourselves with the responses, coded inductively, organized them into

candidate themes, reviewed them for internal coherence, defined them, and subsequently interpreted them in relation to the quantitative findings.

3.0 Results

For our results, we considered the participants who completed the total duration of the training and participated in the initial and final survey. This resulted in a total of 42 participants with a total of 640 sessions, 13100 training blocks, and 42 entries for each of the surveys

3.1 Quantitative

Table 1: N-Back Training Progress Report

Level	N participants	Mean Accuracy (%)	SD (%)
1-back	42	77.2	~3.5
2-back	42	75.4	~4.0
3-back	31	78.2	~2.5
4-back	12	77.9	~5.0
5-back	8	75.5	~6.0
6-back	2	79.6	3.9
7-back	1	76.4	—
8-back	1	78	—

The whole cohort of 43 participants participated in our adaptive Dual N-Back training, wherein they would progress through the levels based on performance in previous levels. The minimum level was 1, which every participant could complete, and the maximum level achieved was 8, which only 1 participant could reach. All of them could participate in the 1st level (mean Accuracy = 77.2%, SD = ~3.5%); these levels progressively got harder and required more working memory capabilities. Hence, we observed a drop in participants as we progressed through the levels. We observed that the drop between level = 3 (N=31) and level = 4 (N=12) was the sharpest. Table 1 explains this relationship.

To check the progress of the whole cohort, we analysed all 13,100 data points to calculate the maximum Dual N- Back level achieved by each participant on the last day of the training (day 15/16) as compared to the first day of the training. We observed that on Day 1 of the training, 24 participants could achieve level 2, 15 could achieve level 3, and 3 of the participants remained at level 1. The data from the final day shows that improvement in test performance did occur. We saw 18 of our participants in level 3, 13 in level 2, 5 in level 5, 3 in level 1, 2 in level 4, and 1 in level 8. These observations suggested a general trend of improvement in the working memory of the participant in Dual N-Back Task performance.

3.2 Qualitative Data

Although Quantitative measures pointed to overall improvement in task performance in Dual N-Back training, our qualitative surveys conducted pre and post-training are necessary to establish the ecological validity of our findings.

In the initial pre-training survey, we had 42 participants aged 9 to 13, enrolled in grades 5 through 8, with the vast majority of them (~90%) reporting no previous experience with structured computerized brain training. We observed that participants were relatively split between "Neutral" and "Very Motivated" stances towards the upcoming training sessions. Following the 15/16 day training protocol, the final survey was conducted. It recorded reflections on task difficulty, perceived benefits, and environmental modulators. Analysis of the final survey responses generated four dominant themes.

a. Theme 1: Enhanced Attentional Control and Distractor Inhibition

We observed a prevalent self-reported improvement in sustained attention and ability to focus after the training. Participants frequently reported perceived gains in concentration regarding academic tasks. One participant stated:

"I am able to complete my homework with 20% fewer distractions"

Another participant mentioned being less influenced by digital distractions:

"I can now finish my homework in one sitting without checking my phone every 10 minutes"

These self-reported improvements are strongly in line with previous literature suggesting that Dual N-Back training enhances inhibitory control.

b. Theme 2: Subjective Expansion of Working Memory Capacity

In relation to their daily academic retention, participants reported noticeable improvements in their ability to retain and manipulate information. A participant recalled:

"I remember things quicker and better"

While another participant observed improvement in class performance:

"I remember formulas and key points from class without having to re-read them multiple times"

These observations directly reflect the near transfer effects of working memory training.

c. Theme 3: Academic Self-Efficacy, Behavioural Flexibility and Improved Reasoning

We observed a notable effect emerge regarding academic confidence and problem-solving. This may be caused by the increasing Working Memory efficiency due to the training. As working memory efficiency increases, cognitive load for the same task is reduced as compared to before, and this frees up resources for other higher-order reasoning tasks.

A participant reported a shift in their confidence:

"I feel more confident while answering questions in class and participating in discussions. I am less hesitant to speak up"

Another reported improved clarity:

"I try different approaches in math problems now instead of getting stuck on the first step"

This is interesting as mathematics is highly reliant on fluid intelligence (Gf).

d. Theme 4: Environmental Modulators of cognitive load

We observed that participants had a high degree of awareness regarding how environmental factors modulated their performance.

A participant cited "quiet, stable, airy and well-lit" environment as a catalyst for high performance. On the other hand, auditory distractors and visual distractors like television, chatter, construction noise, and digital displays were reported as primary disruptors of focus.

A participant stated:

"When my father watches TV and switches to the IPL matches it distracts me"

These reports highlight the key role environmental factors play in influencing executive functioning.

e. Theme 5: Perceptions of Societal Utility

The participants who completed the training were asked about their opinions on how and if such cognitive training could be beneficial for broader sectors of society. They reported an understanding of the spread-out applications of working memory.

A participant stated:

"It would help everyone: kids, teenagers, adults, and even old people. It will help them to exercise their brain daily and improve their mental abilities"

While another suggested that this training could result in increased attentional spans, which could help individuals "understand problems more quickly, which directly increases productivity".

4.0 Discussion

Our study aimed to evaluate the effects of Dual N-Back training over a short period of 15/16 days in school environments. We focused on age groups 9 through 13, to focus on a sample where the possibility of plasticity is high and quick. We synthesized both qualitative and quantitative data into a mixed-methods framework for each of the 42 participants:

Table 2: Merged quantitative performance tracking and qualitative thematic trajectories

Participant	Baseline Disposition	Final Outcome	Day 1 N-Level	Final Day N-Level	Total Days
sub 01	Baseline Ambivalence	Enhanced Attentional Control	1	5	15
sub 02	Baseline Ambivalence	Perceived Lack of Efficacy	2	1	15
sub 03	High Initial Engagement	Enhanced Attentional Control	2	2	15
sub 04	Baseline Ambivalence	Academic Self-Efficacy	1	1	15
sub 05	High Initial Engagement	Working Memory Expansion	3	3	15
sub 06	Baseline Ambivalence	Working Memory Expansion	2	5	15
sub 07	High Initial Engagement	Academic Self-Efficacy	3	2	15

sub 08	Baseline Ambivalence	Enhanced Attentional Control	2	2	15
sub 09	Baseline Ambivalence	Enhanced Attentional Control	3	3	16
sub 10	Baseline Ambivalence	Academic Self-Efficacy	3	3	16
sub 11	Baseline Ambivalence	Enhanced Attentional Control	3	3	16
sub 12	Baseline Ambivalence	Working Memory Expansion	2	2	15
sub 13	Baseline Ambivalence	Enhanced Attentional Control	3	3	16
sub 14	Initial Skepticism	Perceived Lack of Efficacy	2	2	16
sub 15	Baseline Ambivalence	Enhanced Attentional Control	3	3	16
sub 16	Baseline Ambivalence	Working Memory Expansion	2	5	15
sub 17	Baseline Ambivalence	Enhanced Attentional Control	3	3	16
sub 18	Baseline Ambivalence	Working Memory Expansion	3	5	15
sub 19	High Initial Engagement	Academic Self-Efficacy	2	2	15
sub 20	Baseline Ambivalence	Working Memory Expansion	2	2	16
sub 21	Baseline Ambivalence	Enhanced Attentional Control	3	3	16
sub 22	Baseline Ambivalence	Perceived Lack of Efficacy	2	2	16
sub 23	High Initial Engagement	Academic Self-Efficacy	2	5	15
sub 24	High Initial Engagement	Enhanced Attentional Control	1	2	15
sub 25	Baseline Ambivalence	Academic Self-Efficacy	2	2	15
sub 26	Baseline Ambivalence	Academic Self-Efficacy	2	1	15
sub 27	Baseline Ambivalence	Enhanced Attentional Control	3	3	15
sub 28	Baseline Ambivalence	Enhanced Attentional Control	3	8	15
sub 29	Baseline Ambivalence	Working Memory Expansion	2	3	15
sub 30	Baseline Ambivalence	Enhanced Attentional Control	2	3	15
sub 31	Baseline Ambivalence	Enhanced Attentional Control	2	2	15
sub 32	Initial Skepticism	Academic Self-Efficacy	2	3	15
sub 33	High Initial Engagement	Enhanced Attentional Control	2	3	15
sub 34	High Initial Engagement	Working Memory Expansion	2	3	15

sub 35	High Initial Engagement	Working Memory Expansion	2	3	15
sub 36	High Initial Engagement	Academic Self-Efficacy	3	4	15
sub 37	High Initial Engagement	Academic Self-Efficacy	2	2	15
sub 38	Baseline Ambivalence	Academic Self-Efficacy	2	3	15
sub 39	Baseline Ambivalence	General Maintenance	3	3	15
sub 40	Baseline Ambivalence	Academic Self-Efficacy	2	3	15
sub 41	Baseline Ambivalence	Enhanced Attentional Control	2	2	15
sub 42	High Initial Engagement	Academic Self-Efficacy	3	4	15

Our quantitative results demonstrated an overall modest progression in task performance across all participants, with notable individual variability. The sharpest drop in progress was observed between levels 3 and 4. This can be caused by the dual modality demand of our training paradigm, which forces central executive adaptation (Baddeley, 1992). It may be possible that the ability to push beyond this level 3 plateau was observed as an influencing factor of whether the improvements in working memory generalize to unpracticed domains.

Although the current literature is conflicted regarding the extent of far transfer of training effects to generalized intelligence and global academic outcomes (Melby-Lervåg & Hulme, 2013; Soveri et al., 2017), the qualitative survey responses indicate real-world behavioural adaptations. In our qualitative dataset, themes 1: Enhanced Attentional Control and Distractor Inhibition and theme 4: Environmental Modulators of cognitive load highlight self-reported improvements in distraction inhibition and sustained attention, which can act as a counterbalance to the attentional fragmentation caused by modern trends of multi-media consumption (Cain et al., 2016; Uncapher et al., 2017). Participants reported frequently that they can now focus better on homework and filter out noise like television and phones. This filtering out of irrelevant information is a key skill that influences performance in Dual N-Back training.

Theme 3 of our qualitative analysis, Academic Self-Efficacy, Behavioural Flexibility, and Improved Reasoning, shows us how a structured cognitive training like Dual N-Back can improve our psychological scaffolding through confidence gained from mastering such a high-friction adaptive leveling task. In our sample, this empowered students who previously got stuck with mathematics to approach similar problems in multiple ways, and to use new problem-solving methods that rely on fluid intelligence. (Jaeggi et al., 2008).

A certain subset of participants whose performance in the task plateaued at level 2 or 3 still reported gains in memory and academic efficacy. This may be a result of increased metacognitive awareness leading to independent learning styles and strategies. Even when measurable outcomes like N-Back level or accuracy may plateau over time (Shinomiya et al., 2015), the training programme still requires the participants to be engaged repeatedly with increasingly demanding working memory tasks, where they have to monitor and regulate their

cognitive processes. This may have increased self-awareness and nudged them into forming learning strategies to improve performance. Hence, we may say that this training may also be used to support the development of adaptive cognitive and behavioural skills associated with learning.

5.0 Conclusion

Our study highlights the potential value of adaptive Dual N-Back training as a school-based cognitive intervention. Our quantitative findings demonstrated overall improvements in task performance across the training period, although progression in difficulty levels plateaued for many participants at the 3-back level. Although this indicates measurable gains in task performance may stabilize over a relatively short intervention period, our qualitative findings revealed meaningful perceived improvements in participants' cognitive functioning and academic experience, and day-to-day activities.

Interestingly, several participants whose quantitative performance plateaued continued to report subjective gains in memory, attention, and academic confidence. This divergence between objective performance and perceived improvement suggests that the benefits of the intervention may apply beyond Dual N-Back scores alone. The constant need to monitor, update, and manipulate information may have led the students to become more reflective about their cognitive processes, promoting independent learning strategies even when measurable performance gains began to stabilize (Shinomiya et al., 2015).

Taken together, the quantitative and qualitative findings suggest that adaptive Dual N-Back training may serve as a valuable complement to existing educational practices. Although further randomized controlled studies with larger and more diverse samples are required to establish long-term efficacy and transfer to academic achievement, the present findings indicate that structured working memory training has the potential to enhance attention, metacognitive awareness, and learner confidence.

Works Cited

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 8, pp. 47–89). Academic Press.

Cain, M. S., Leonard, J. A., Gabrieli, J. D. E., & Finn, A. S. (2016). Media multitasking in adolescence. Psychonomic Bulletin & Review, 23(6), 1932–1941. https://doi.org/10.3758/S13423-016-1036-3

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. Proceedings of the National Academy of Sciences, 105(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. Developmental Psychology, 49(2), 270–291.

Shinomiya, K., et al. (2015). Developmental trajectories for attention and working memory in healthy school-aged children. Brain and Development, 37(8), 789–796.

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of N-back training studies. Psychonomic Bulletin & Review, 24(4), 1077–1096.

Uncapher, M. R., Wagner, A. D., & colleagues. (2017). Media multitasking and cognitive, psychological, neural, and learning differences. Pediatrics, 140(Supplement 2), S62–S66.

Can Poetry Cultivate Empathy? A Review of Psychological Mechanisms and Empirical Evidence By Penny Wei

Abstract

This literature review examines how poetry may foster empathy through perspective-taking, emotional simulation, imagery, and ambiguity. It synthesizes research on literary reading, poetry-specific emotional responses, educational applications, and prosocial behavior. While evidence suggests that poetry can facilitate empathic engagement, particularly by encouraging readers to imagine unfamiliar perspectives and emotions, research on lasting changes in empathy remains limited. The review argues that poetry is best understood as a space for practicing the imagination of other minds rather than a direct or universal mechanism for producing empathy

Keywords poetry, empathy, literary-reading, fiction, perspective-taking, emotional simulation, imagery, Theory of Mind

Introduction

Empathy is the ability to understand and respond to another person's emotional experience. It allows individuals to move beyond their own perspectives and engage with experiences that may be unfamiliar or distant from their own. Because literature invites readers to encounter other voices, emotions, and ways of seeing the world, it has long been considered a potential medium for developing this capacity.

Poetry is particularly compelling in this regard. Unlike ordinary communication, poetry often conveys experience indirectly through metaphor, imagery, rhythm, ambiguity, and subjective voice. Readers must actively interpret what a speaker feels, reconstruct experiences that may not be explicitly described, and imagine perspectives different from their own. In doing so, poetry may transform another person's experience from something merely observed into something mentally and emotionally experienced.

Yet the relationship between poetry and empathy is not necessarily straightforward. Feeling an emotion while reading a poem does not automatically mean that a reader has become more capable of understanding or responding to others. The extent to which poetic experiences translate into lasting empathic development therefore remains an important question.

This literature review asks: How does poetry foster empathy, and what psychological mechanisms explain its potential empathic effects? It examines the roles of perspective-taking, emotional simulation, imagery, and ambiguity, before considering empirical evidence, educational applications, and the relationship between literary empathy and prosocial behavior. Ultimately, it argues that poetry is best understood not as an automatic source of empathy, but as a space in which readers can practice imagining and inhabiting other minds.

Defining Empathy

Empathy is the complex capability of an individual to sense, feel and understand the emotions of others, along with the ability to imagine what someone else might be thinking or feeling. Empathy has been viewed as an important aspect of human social functioning, influencing interpersonal relationships and moral judgment. Within psychological research, empathy is often conceptualized as a multidimensional construct consisting of both cognitive empathy and affective empathy. Cognitive empathy refers to the ability to recognize, interpret, and understand another person's internal states, often described as perspective-taking or theory of mind. Affective empathy, in contrast, involves the capacity to share or resonate with another person's emotional experiences, such as experiencing sadness in response to another individual's suffering. Researchers have also identified a third dimension, compassionate empathy, which extends beyond understanding and emotional resonance to include motivation to provide support or alleviate another person's distress. These components work together to shape how individuals perceive and respond to the experiences of others. While empathy was historically studied primarily within the context of face-to-face social interactions, contemporary psychological research has expanded its scope to examine how symbolic forms of communication, including literature and poetry, can evoke empathic responses even toward individuals who are distant in time, culture, or personal experience. Poetry provides a unique medium for cultivating empathy because it communicates subjective emotional experiences. Unlike factual descriptions of events, poetry often presents the internal world of an individual through imagery, metaphor, rhythm, and first-person perspectives. These literary features may allow readers to temporarily inhabit another person's emotional reality, facilitating processes such as perspective-taking and emotional stimulation. Therefore, understanding the relationship between poetry and empathy requires examining the psychological mechanisms through which poetic engagement influences readers' perceptions, emotions, and social cognition.

Literature and Empathy

Literature has long been theorized as a medium capable of expanding human understanding by allowing individuals to encounter perspectives and experiences beyond their immediate social environments. Unlike direct interpersonal interactions, literary texts provide access to characters' internal states, motivations, emotions and conflicts, creating opportunities for readers to engage in forms of psychological perspective-taking. Contemporary psychological research suggests that literary engagement may strengthen empathic abilities by exercising the cognitive and emotional processes involved in understanding others. Mar and Oatley (2008) proposed that fiction functions as a form of social simulation, allowing readers to mentally rehearse interpersonal situations without directly experiencing them. Through engagement with fictional characters, readers practice interpreting intentions and social relationships, processes that overlap with the mechanisms underlying empathy. Empirical evidence supporting the relationship between literary reading and empathic cognition was provided by Kidd and Castano (2013), who found that brief exposure to literary fiction improved participants' performance on measures of theory of mind, the ability to infer and understand others' mental states. The authors

argued that literary fiction requires readers to interpret complex characters whose motivations are often ambiguous, encouraging deeper perspective-taking. Unlike popular narratives that often provide clear moral judgments and predictable character motivations, literary fiction frequently presents psychologically complex individuals whose intentions must be actively inferred by readers. This interpretive process may strengthen cognitive empathy by requiring readers to construct mental representations of perspectives different from their own. Neuroscientific research further suggests that literary engagement recruits neural systems involved in social cognition. Tamir et al. (2016) found that reading fiction was associated with activity in the brain's default network, a system involved in imagining others' thoughts, feelings, and perspectives. These findings support the idea that literary reading is a form of mental simulation that engages processes relevant to empathy. Beyond cognitive perspective-taking, literature may cultivate affective empathy through emotional transportation and identification. Oatley (1999) argued that readers often simulate characters' emotional experiences, creating a psychological connection between the reader and fictional individuals. Through this process, readers may experience and fully understand emotions that correspond with, rather than simply observe, another person's circumstances, cultivating sympathy.

Although research on literature and empathy has primarily focused on narrative fiction, poetry represents a distinct form of literary expression that may produce similar or even intensified empathic effects. Whereas novels often develop empathy through extended character development and narrative immersion, poetry frequently compresses emotional experiences into concentrated linguistic forms. Through multiple mechanisms such as first-person voice, metaphor, imagery, and emotional ambiguity, poetry may provide readers with direct access to subjective experiences, creating unique pathways for empathic engagement.

Psychological Mechanisms of Poetry

Perspective Taking

One of the primary psychological mechanisms through which poetry may cultivate empathy is its capacity to facilitate perspective-taking. Perspective-taking involves the ability to temporarily adopt another individual's viewpoint and understand their emotional and cognitive experiences. Because poetry frequently presents subjective experiences through first-person narration and intimate reflection, readers are encouraged to dissect and interpret the speaker's internal world rather than simply observe external events. Research on narrative empathy suggests that engaging with texts from another person's perspective activates processes involved in understanding others' mental states. Mar and Oatley (2008) argue that literary experiences function as "simulations" of social interactions, allowing readers to practice interpreting emotions, intentions, and beliefs. Poetry may intensify this process because it often removes narrative distance between the reader and speaker. Through direct address and personal voice, poems position readers as witnesses to another person's emotional reality. For example, a poem written from the perspective of a refugee, a grieving individual, or someone experiencing

discrimination invites readers to imaginatively inhabit that experience rather than just providing information about suffering. This process aligns with cognitive empathy by requiring readers to construct perspectives that may differ from their own. Keen (2007), in *Empathy and the Novel*, argues that literary texts encourage readers to engage in "empathetic imagination," where readers temporarily enter another person's consciousness through narrative representation. Although Keen focuses primarily on fiction, this concept applies particularly strongly to poetry because poems often rely on emotional immediacy and subjective voice. Keen states how "empathy is a vicarious, spontaneous sharing of affect" (Keen, 2007). This emotional sharing distinguishes poetry from informational texts. Through taking on the poet's personal perspective through poetry, readers may experience an emotional response connected to that suffering.

Metaphors

Another mechanism through which poetry may cultivate empathy is metaphor. Unlike literal language, metaphor requires readers to map relationships between seemingly unrelated concepts, encouraging deeper interpretation and emotional understanding. Because emotions are often abstract and difficult to directly describe, poets frequently rely on metaphor to represent internal experiences. Lakoff and Johnson (1980) argue in *Metaphors We Live By* that metaphor is a fundamental literary device through which humans conceptualize abstract experiences. Emotional states such as grief, love, fear, and loneliness are frequently understood through embodied metaphors. For example, describing sadness as "a heavy weight" connects an invisible emotional state to a physical sensation that readers can imagine. Lakoff and Johnson state how "the essence of metaphor is understanding and experiencing one kind of thing in terms of another." Through metaphorical language, poetry allows readers to access emotional experiences that they may not personally share. A reader who has never experienced displacement, loss, or trauma may nevertheless understand these experiences through metaphorical associations that connect unfamiliar emotions with familiar bodily or sensory experiences. Research in cognitive linguistics supports this interpretation. Gibbs (2006) suggests that metaphor comprehension involves embodied simulation, meaning that readers partially recreate sensory and emotional experiences associated with metaphorical expressions. Therefore, poetic metaphors may contribute to empathy by transforming another person's internal state into an experience that feels psychologically accessible.

Ambiguity

Poetry's frequent use of ambiguity may also contribute to empathic development by requiring readers to actively interpret multiple possible meanings. Unlike explicit forms of communication, poems often leave emotions and experiences partially unresolved and not thoroughly explained. This uncertainty requires readers to engage in perspective-taking and consider interpretations beyond their immediate assumptions. Kidd and Castano (2013) demonstrated that literary fiction improves theory of mind because it often contains psychologically complex characters whose thoughts and motivations are not directly explained.

Poetry similarly relies on interpretive openness, requiring readers to infer emotional states from symbols, images, and fragmented narratives.

Ambiguity may therefore function as an exercise in tolerating psychological complexity. When readers encounter a poem with conflicting emotions or indefinite meaning, they must resist making immediate judgments and instead explore alternative perspectives. This process mirrors the cognitive flexibility required for empathy in real-world social interactions. Oatley (1999) suggests that literary reading involves a "meeting of minds," where readers interact with representations of consciousness different from their own. Poetry's compressed and ambiguous structure may intensify this interaction because readers must actively participate in constructing meaning.

Imagery

Poetry may further evoke empathy through its ability to generate vivid mental imagery. Unlike abstract descriptions of emotions, poetry frequently uses sensory language, imagery, and figurative expression to encourage readers to mentally visualize experiences. This imaginative process can create emotional simulations that resemble direct perception. Research in cognitive psychology suggests that mental imagery plays an important role in emotional processing. When individuals imagine another person's experiences, overlapping cognitive and neural processes involved in personal emotional experiences may become activated. By creating sensory representations of another person's world, poetry may allow readers to experience emotional states from an imagined perspective. Tamir et al. (2016) found that fiction reading engages neural networks associated with understanding other people's thoughts and experiences. Similarly, poetry's reliance on imagery may encourage readers to construct detailed internal representations of another person's emotions and circumstances. For example, a poem describing war through images of abandoned homes, empty streets and a child's fear creates an imagined environment through which readers can emotionally engage with experiences distant from their own. As Oatley (2012) explains, literature allows individuals to "simulate" experiences beyond their personal lives. Poetry's concentrated use of imagery may make this simulation particularly powerful by compressing complex emotions into memorable sensory moments.

Poetry's Empathic Effects

The psychological mechanisms discussed above suggest several pathways through which poetry might cultivate empathy. Perspective-taking may allow readers to reconstruct another person's internal state; metaphor and imagery may make unfamiliar emotions psychologically accessible; and ambiguity may require readers to tolerate perspectives that cannot be reduced to a single interpretation. However, identifying plausible mechanisms does not establish that poetry actually produces empathic effects. Empirical research is therefore necessary to determine whether literary engagement is associated with measurable changes in perspective-taking, emotional resonance, empathic understanding, or prosocial behavior. Existing evidence provides preliminary support for these relationships, although the strength of the evidence varies

considerably. Research on literary fiction offers the strongest broader foundation, while a smaller but growing body of research specifically examining poetry suggests that poetry can evoke empathic and emotional responses. At the same time, methodological limitations make it difficult to conclude that poetry directly produces durable increases in empathy.

Literary Reading

Because empirical research specifically examining poetry and empathy remains relatively limited, research on literary reading provides an important theoretical and empirical foundation. Literary fiction is particularly relevant because it requires readers to engage with characters whose thoughts, motivations, and emotional states may differ from their own. This process overlaps closely with cognitive empathy, which involves recognizing and understanding another person's internal state.

One of the most influential experimental studies in this area was conducted by Kidd and Castano (2013). Across five experiments, the researchers examined whether reading literary fiction affected participants' performance on measures of Theory of Mind (ToM), which refers to the ability to infer and understand the mental states of other people. Participants who read literary fiction subsequently performed better on measures of both affective and cognitive Theory of Mind than participants who read nonfiction, popular fiction, or nothing at all. The authors therefore concluded that literary fiction can temporarily enhance Theory of Mind and suggested that engagement with art may influence social-cognitive processes.

These findings are relevant to empathy because Theory of Mind and cognitive empathy involve overlapping processes of perspective-taking. A reader attempting to determine why a character behaves in a particular way must infer beliefs, intentions, emotions, and motivations that are not necessarily stated explicitly. In this sense, literary reading may provide a form of cognitive practice: rather than simply receiving information about another person, readers must construct an internal representation of that person's psychological perspective. This interpretation is consistent with the social-simulation account of literary reading, which proposes that fiction allows readers to mentally simulate interpersonal situations and rehearse social cognition without directly experiencing the situation themselves.

Importantly, however, Theory of Mind should not be treated as synonymous with empathy. A person may accurately understand another individual's emotions without emotionally sharing those emotions or feeling motivated to help. Kidd and Castano's findings therefore provide evidence for a cognitive component underlying empathy rather than proof that literary fiction produces comprehensive or durable empathic development.

Furthermore, the original findings have been subject to replication debates. Panero et al. (2016), for example, combined data from three independent research groups and did not find significantly higher performance on the Reading the Mind in the Eyes Test after participants read literary fiction compared with other conditions. Their results therefore failed to support a short-term causal effect of literary fiction on Theory of Mind. At the same time, the researchers found that lifetime exposure to fiction was consistently associated with Theory of Mind

performance, leaving open the possibility that longer-term reading habits may relate to social cognition even if a single reading session does not immediately improve it.

Kidd and Castano later conducted three additional replication studies. Their analysis concluded that one of the three replications successfully reproduced the earlier effect, while the other two were considered uninformative; they also noted that previous close replications and correlational evidence supported an association between literary-fiction experience and Theory of Mind. Thus, the broader literature does not provide an uncomplicated causal relationship between literary reading and empathy. Instead, it suggests that literary engagement may activate or exercise social-cognitive processes, while the magnitude, reliability, and durability of these effects remain contested.

This distinction is particularly important when extending literary research to poetry. If literary fiction can engage perspective-taking, poetry may activate similar processes through its own formal characteristics. Yet poetry cannot simply be assumed to produce the same effects as narrative fiction. Unlike novels, poems often contain little character development or explicit narrative structure. Their potential empathic effects may instead depend more heavily on emotional voice, imagery, metaphor, rhythm, and the reader's active construction of meaning.

Poetry-Specific Impact on Empathy

Direct empirical research on poetry provides a more specific test of whether poetic reading generates emotional and empathic responses. Although this body of research remains smaller than the literature on narrative fiction, recent studies suggest that empathy and mental imagery are important components of the emotional experience of reading poetry.

Pițur and Miu (2025) provide particularly relevant evidence. Their study investigated how reading experience and individual psychological characteristics—including empathy, visual imagery, movement imagery, auditory imagery, and fantasy proneness—relate to emotions elicited by poetry. Participants first completed measures of their reading experience and psychological traits and then participated in a poetry-reading session. They subsequently reported the emotional intensity they experienced while reading and the psychological processes involved in their reading experience.

The researchers found evidence that empathy and imagery experienced during reading helped explain the relationship between previous poetry-reading experience and emotional intensity. In other words, more experienced readers did not simply report stronger emotions independently of psychological processes; the relationship between reading experience and emotional response was mediated by processes including empathy and imagery.

This finding is significant because it connects two mechanisms discussed earlier in this review—empathy and imagery—with an observable response to poetry. The results suggest that poetry may not produce emotional responses merely because its subject matter is emotionally charged. Rather, readers' psychological engagement with the poem appears to matter. When readers empathize with authors or characters and generate vivid mental imagery, their emotional response to the poem may become more intense.

More recent work by Pițur, Tufar, and Miu (2025) further emphasizes the role of imagery in poetry-elicited emotion. In a study involving hard-of-hearing participants and controls, participants read poetry for 30 minutes and subsequently rated their emotional responses as well as their experiences of empathy and different forms of mental imagery. The researchers focused particularly on auditory imagery—the experience of internally "hearing" words while silently reading. They found evidence linking auditory imagery to emotional arousal during poetry reading and examined whether auditory imagery could help explain differences in emotional response.

This research strengthens the argument developed earlier concerning imagery. Poetry may transform abstract emotional states into quasi-perceptual experiences: readers do not merely understand that a poem describes sadness, intimacy, fear, or grief, but may internally simulate aspects of the experience through sensory and emotional imagery. Importantly, however, these studies primarily measure emotional response to poetry, not whether readers become more empathic people after repeated exposure to poetry. Their findings therefore support the existence of empathic and emotional processes during reading without establishing long-term changes in trait empathy.

A related line of research is provided by Lüdtke, Meyer-Sickendieck, and Jacobs (2014), who developed and tested the mood empathy hypothesis of poetry reception. The hypothesis proposes that poems depicting the moods of people, situations, or objects can lead readers to mentally simulate and effectively resonate with the depicted state. In this model, the reader's interaction with the poem produces something closer to emotional participation than detached observation: the reader may experience the mood represented by the poem or a closely associated emotional state.

Their research is particularly relevant to affective empathy. Whereas Theory of Mind research primarily concerns understanding another person's mental state, the mood empathy hypothesis proposes that readers can experience an emotional resonance with the state represented in a poem. Their findings provided evidence consistent with the hypothesis, showing that reading poetry can produce both aesthetic responses and nonaesthetic emotional responses associated with the depicted situations and moods.

Jacobs et al. (2016) extended this research using a multimethod approach that combined qualitative and quantitative descriptions of poems with subjective, behavioral, and peripheral-physiological measures. Their study tested the mood empathy hypothesis as part of the Neurocognitive Poetics Model and reported results consistent with the hypothesis that poems depicting moods engage empathic and immersive processes.

These studies suggest that poetry may facilitate what can be understood as emotional simulation. A poem can provide enough information for readers to construct a mental representation of a situation, while its imagery, language, and formal features encourage readers to experience aspects of the emotional state represented within it. This provides empirical support for the theoretical mechanisms proposed earlier: poetry may cultivate empathic

engagement not simply by telling readers about another person's experience, but by encouraging them to mentally and emotionally simulate it.

Nevertheless, the distinction between empathic response and empathic development remains essential. The fact that a reader experiences empathy while reading a poem does not necessarily mean that the experience will produce a lasting increase in the person's general capacity for empathy. Current poetry research provides stronger evidence for the former than the latter.

Poetry's Impact Under Educational and Applied Settings

Evidence from educational settings provides another perspective on poetry's potential to cultivate empathy. Rather than examining poetry only as an aesthetic object, these studies deliberately use poetry as a pedagogical tool to help students encounter experiences and perspectives beyond their own.

Hatem and Ferrara (2001), for example, examined a creative-writing elective designed for preclinical medical students. The course sought to strengthen students' writing, observation, and reflective abilities in relation to their medical education. Through qualitative analysis of students' writing, the researchers identified themes related to the development of more humane approaches to medical practice.

This study is useful because medical education provides a context in which empathy has a concrete social function. Physicians must understand patients not merely as collections of symptoms but as individuals with fears, histories, relationships, and subjective experiences. Creative writing can encourage students to slow down, observe closely, and articulate experiences that are difficult to express through clinical language. However, because Hatem and Ferrara examined creative writing broadly rather than isolating poetry as an independent intervention, their findings should be interpreted as evidence for the broader reflective potential of literary and creative writing rather than as proof that poetry alone increases empathy.

More direct evidence comes from Mood's (2018) study of beginning nursing students. Eight nursing students participated in a seminar in which they read a poem and discussed it through guided questions concerning empathy and professional identity. Following the seminar, students wrote reflections about their most significant learning. Through discussion and reflection, students developed new understandings of empathy and incorporated those ideas into their imagined professional identities. Six of the eight students identified the poetry seminar as their most important learning experience of the week.

Although the study's sample was extremely small, its findings illustrate an important function of poetry: it can make an abstract concept such as empathy concrete and discussable. Rather than defining empathy theoretically, students encountered an artistic representation of human experience and were required to interpret it, discuss it with others, and relate it to their own future professional roles. Poetry therefore served as a mediating object through which students could practice articulating another person's perspective.

Vicino's (2022) action-research study provides perhaps the most directly relevant educational evidence for poetry and social empathy. Over six weeks, six secondary-school students participated in a social-justice poetry unit combining close reading and reader-response approaches. The researcher measured empathy before and after the unit through Likert scales and complemented these quantitative measures with surveys, interviews, observations, and student artifacts. The study found that students' empathy levels were elevated following the unit. Qualitative evidence suggested that social-justice poetry encouraged students to consider perspectives outside their everyday lived experiences and increased their understanding of people different from themselves.

The significance of this study lies not only in its reported increase in empathy but also in the type of poetry used. Social-justice poetry explicitly foregrounds experiences involving inequality, discrimination, identity, and marginalization. Such poetry may therefore create opportunities for perspective-taking by placing readers in contact with experiences that are socially or personally distant from their own. In this sense, poetry can function as a bridge between aesthetic experience and social understanding.

However, Vicino's findings should be interpreted cautiously. The study involved only six students and was conducted within a single classroom, making it impossible to determine whether the observed increase would generalize to larger populations. The absence of a large control group also makes it difficult to establish that poetry itself, rather than classroom discussion, teacher interaction, or repeated reflection, caused the change. Nevertheless, the study offers valuable preliminary evidence that poetry can be intentionally incorporated into educational environments to encourage perspective-taking and empathic thinking.

Taken together, these educational studies suggest that poetry may be most effective when it is accompanied by reflection and discussion. Reading alone may initiate an emotional or imaginative response, while guided interpretation can encourage readers to consciously connect that response to another person's experience. Poetry's empathic potential may therefore arise not from the text alone but from an interaction among text, reader, and social context.

Prosocial Behavior

An important question is whether empathic responses generated through literary reading extend beyond the reading experience itself. Empathy is not only the ability to understand or share another person's emotional state; it can also motivate compassionate or prosocial behavior. If literature cultivates empathy but has no effect on how individuals respond to others, its social consequences may be more limited than often assumed.

Koopman and Hakemulder (2015) reviewed empirical research concerning the relationship between literary reading, empathy, self-reflection, and potential prosocial outcomes. They note that scholars have frequently proposed a chain in which literary reading generates empathic responses and self-reflection, which may subsequently contribute to more prosocial behavior. However, they emphasize that the empirical evidence for this complete causal chain remains limited.

Koopman's empirical work is particularly important because it examines the conditions under which literary reading produces empathic reactions rather than assuming that all literary reading is equally effective. In a 2015 study, participants read different versions of a literary narrative concerning the loss of a child. The study investigated how textual "literariness," emotional response, trait empathy, previous personal experience, and previous literary exposure related to empathy and reflection after reading.

The findings indicate that readers' responses cannot be explained solely by the literary text itself. Individual characteristics—including existing empathy and previous experiences—also influenced readers' responses. This supports a more interactive understanding of literary empathy: a text may provide opportunities for empathic engagement, but readers bring their own emotional capacities, memories, and experiences to the reading process.

Research connecting literary reading to prosocial behavior provides further evidence that emotional engagement can sometimes extend beyond the text. For example, research summarized in the literary-empathy literature has found that transportation into narratives can be associated with empathy and prosocial outcomes, while generating imagery while reading fiction has been shown to increase empathy and prosocial behavior. These findings are important because they suggest a possible pathway from imaginative simulation to behavioral response, rather than treating empathy as an isolated emotional experience.

However, the distinction between feeling empathy and acting on empathy remains critical. A reader may feel sadness for a fictional character without changing their behavior toward real people. Likewise, understanding the suffering of a marginalized group does not automatically result in charitable action, political engagement, or interpersonal compassion. Koopman's broader review emphasizes precisely this uncertainty: although theoretical accounts frequently connect literary empathy with prosocial behavior, researchers are still far from establishing when and how literary reading produces durable empathic effects in real-world relationships.

Thus, prosocial behavior should be considered a possible downstream outcome rather than an established consequence of poetry reading. The evidence currently supports a more cautious model:

Poetry and literature → Emotional Engagement and Perspective-taking → Empathic Response → Possible Prosocial Motivation → Possible Behavior

Each arrow represents a separate psychological process, and the evidence is considerably stronger for the earlier stages than for the final behavioral outcome.

Limitations

Despite promising findings, the empirical literature does not justify the simple conclusion that poetry necessarily increases empathy. Several methodological limitations complicate this interpretation.

First, many poetry-specific studies involve small samples. Mood's nursing study included only eight students, while Vicino's social-justice poetry study included six. Such samples can provide detailed qualitative evidence about how participants experience poetry, but they have limited statistical power and cannot establish whether similar effects would appear across different populations.

Second, poetry is often studied alongside other forms of literature or creative writing. Hatem and Ferrara's study, for example, examined a creative-writing course rather than isolating poetry. Consequently, positive findings from literary or creative-writing interventions cannot automatically be attributed to poetry itself.

Third, many studies rely partly on self-report measures. Participants may report feeling more empathetic, emotionally engaged, or reflective after reading a poem, but self-reported empathy does not necessarily correspond to observable behavior. This distinction is especially important because empathy contains multiple dimensions. A reader may report understanding another person's feelings without demonstrating greater perspective-taking in a behavioral task or greater willingness to help that person.

Fourth, researchers frequently measure short-term emotional responses rather than long-term empathic development. Pițur and Miu's work, for example, provides evidence that empathy and imagery experienced during poetry reading are associated with the intensity of emotional responses. This is strong evidence that empathic processes can occur during poetry reception, but it does not establish that repeated poetry reading produces a stable increase in trait empathy

Fifth, existing empathy may influence how readers respond to poetry rather than poetry necessarily creating empathy. Pițur, Tufar, and Miu measured both trait empathy and empathy experienced during reading and found meaningful relationships between readers' dispositional characteristics and their psychological processes during poetry reception. This raises an important causal question: do empathic people experience poetry more intensely, does poetry strengthen empathy, or do both processes influence one another?

The broader literary-fiction literature illustrates this problem particularly clearly. Although Kidd and Castano (2013) initially reported that literary fiction temporarily improved Theory of Mind, subsequent research failed to consistently reproduce the short-term causal effect. Panero et al. (2016) found no significant improvement in Theory of Mind after reading literary fiction, while later work produced a more mixed replication record. These conflicting results caution against interpreting a single reading session as sufficient to produce measurable changes in social cognition.

Finally, evidence of emotional resonance should not be equated with evidence of moral transformation. Poetry may cause readers to experience sadness, identification, sympathy, or emotional resonance without necessarily changing their attitudes or behavior. Even research supporting relationships between literary engagement and empathy emphasizes that the pathway from reading to social behavior is complex and remains incompletely understood. Koopman

argues that researchers are still “quite a long way” from determining when and how literary reading affects empathy for others.

These limitations do not invalidate the hypothesis that poetry can cultivate empathy. Instead, they suggest that the relationship should be understood as conditional and process-based rather than automatic and causal. Poetry appears capable of creating opportunities for empathic engagement, particularly through perspective-taking, emotional simulation, imagery, and identification. Whether these momentary experiences accumulate into lasting changes in empathic ability depends on factors including the reader's existing characteristics, the characteristics of the text, the social context in which it is read, and opportunities for subsequent reflection and action.

Conclusion

Taken together, the existing evidence supports a cautious but meaningful relationship between poetry and empathy. Research on literary fiction provides evidence that literary engagement can activate processes related to Theory of Mind and perspective-taking, although the causal strength of these effects remains debated. Poetry-specific research provides more direct evidence that readers experience empathy, imagery, emotional simulation, and mood resonance during poetic reading. Educational studies further suggest that poetry can be deliberately used to encourage students to consider experiences outside their own, particularly when reading is accompanied by discussion and reflection.

The strongest conclusion supported by the evidence is therefore not that poetry automatically makes people more empathetic, but that poetry can create conditions conducive to empathic engagement. Its compressed language, subjective voice, metaphor, ambiguity, and sensory imagery can require readers to reconstruct experiences that may differ substantially from their own. During this process, readers may understand another perspective cognitively, simulate aspects of another person's emotional state, and experience affective resonance with that state. The distinction between empathic experience and empathic development nevertheless represents an important gap in the literature. Existing research provides relatively strong evidence that empathy can occur *during* literary and poetic engagement, but substantially less evidence demonstrates that these experiences produce durable changes in individuals' general capacity for empathy. Future research should therefore move beyond immediate self-reported emotional responses and examine whether repeated engagement with poetry produces measurable changes in perspective-taking, empathic concern, and prosocial behavior over time.

In this sense, poetry may be best understood not as a direct “training program” for empathy, but as a psychological space for practicing the imagination of other minds. Its potential lies in repeatedly asking readers to cross the boundary between their own experience and someone else's—to inhabit an unfamiliar perspective, tolerate ambiguity, simulate an unfamiliar emotional state, and ultimately decide how that understanding might shape their response to others.

Works Cited

Gibbs, R. W. (2007). Embodiment and cognitive science. Cambridge University Press.

Harenski, C. L., et al. (2017). Socio-neuro risk factors for suicidal behavior in criminal offenders with psychotic disorders. Social Cognitive and Affective Neuroscience, 12(1), 70–80. https://doi.org/10.1093/scan/nsw164

Hatem, D., & Ferrara, E. (2001). Becoming a doctor: Fostering humane caregivers through creative writing. Patient Education and Counseling, 45(1), 13–22. https://doi.org/10.1016/S0738-3991(01)00135-5

Keen, S. (2007). Empathy and the novel. Oxford University Press.

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. Science, 342(6156), 377–380. https://doi.org/10.1126/science.1239918

Koopman, E. M. E. (2015). Empathic reactions after reading: The role of genre, personal factors and affective responses. Poetics, 50, 62–79. https://doi.org/10.1016/j.poetic.2015.02.008

Koopman, E. M. E. (2016). Effects of "literariness" on emotions and on empathy and reflection after reading. Psychology of Aesthetics, Creativity, and the Arts, 10(1), 82–98. https://doi.org/10.1037/aca0000041

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. University of Chicago Press.

Lüdtke, J., et al. (2014). Immersing in the stillness of an early morning: Testing the mood empathy hypothesis of poetry reception. Psychology of Aesthetics, Creativity, and the Arts, 8(3), 363–377. https://doi.org/10.1037/a0036826

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. Perspectives on Psychological Science, 3(3), 173–192. https://doi.org/10.1111/j.1745-6924.2008.00073.x

Mood, L. (2018). Poetry: An innovative teaching strategy for exploring empathy with beginning nursing students. Journal of Nursing Education, 57(5), 315–318. https://doi.org/10.3928/01484834-20180420-13

Oatley, K. (1999). Meetings of minds: Dialogue, sympathy, and identification, in reading fiction. Poetics, 26(5–6), 439–454. https://doi.org/10.1016/S0304-422X(99)00011-X

Oatley, K. (2012). The cognitive science of fiction. Wiley Interdisciplinary Reviews: Cognitive Science, 3(4), 425–430. https://doi.org/10.1002/wcs.1185

Panero, M. E., et al. (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. Journal of Personality and Social Psychology, 111(5), e46–e54. https://doi.org/10.1037/pspa0000064

Pițur, S., & Miu, A. C. (2025). Poetry-elicited emotions: Reading experience and psychological mechanisms. Psychology of Aesthetics, Creativity, and the Arts, 19(2). https://doi.org/10.1037/aca0000525

Pițur, S., et al. (2025). Auditory imagery and poetry-elicited emotions: A study on the hard of hearing. Frontiers in Psychology, 16. https://doi.org/10.3389/fpsyg.2025.1509793

Tamir, D. I., et al. (2016). Reading fiction and reading minds: The role of simulation in the default network. Social Cognitive and Affective Neuroscience, 11(2), 215–224. https://doi.org/10.1093/scan/nsv114

Vicino, P. (2022). A finding place: Social justice poetry toward empathy in the secondary English language arts classroom [Doctoral dissertation, University of South Carolina]. Scholar Commons. https://scholarcommons.sc.edu/etd/6949/

The significance of George Balanchine's *Agon:* A groundbreaking neoclassical creation of push and pull By Sonia Schultes

The ballet *Agon* by George Balanchine was groundbreaking for its time. George Balanchine (1904-1983) was a celebrated choreographer of classical ballets and later neoclassical ballets as well. He was also the founder of the New York City Ballet company along with business partner Lincoln Kirsten, as well as the founder of the School for American Ballet (Garis, 1). Igor Stravinsky (1882-1971) was a composer of classical music best known for his works *The Firebird* and *The Rite of Spring*. In order to create the ballet *Agon*, the New York City Ballet received a grant from the Rockefeller Foundation that was specifically meant to advance American music and dance (Feeney). In the 1950s, there were not many ballet schools. The American Ballet Theatre (ABT) company and school was founded in 1939, but it was the only other ballet company and school besides Balanchine's School for American Ballet (SAB) in New York City (Garis, 1). Ballet rose to prominence in America after the hit of Balanchine's *Nutcracker* in 1953 and *The Firebird* in 1949. The style of modern dance, mainly popularized by modern choreographer Martha Graham, was gaining traction alongside ballet with the rise of modernism in the United States a few decades prior to *Agon*'s creation (Barnes, 55). The choreography of *Agon* reflects the tension of the historical period in which it was made, in particular in relation to themes of modernism, 1960s era American race relations, and the exodus of Russian artists after the Russian Revolution (1917-1923). Due to the connection between Stravinsky's unconventional music and Balanchine's modern-inspired choreography, *Agon* is a unique piece that only could have arisen through their connection.

A notable influence on Balanchine's *Agon* from the time of the 1950s was the tension between the U.S. and Russia, escalated by factors including the Space Race and the Cold War (1947-1991). Both Balanchine and Stravinsky, the choreographer and composer of *Agon,* had formerly grown up in Russia before leaving permanently to live as expatriates in the U.S. The conflict between the U.S. and the Soviet Union, often referred to as the Cold War, originated when Russia wanted to expand their sphere of interest while the U.S. defended any country or group of people targeted by Russia (Schlesinger, 30). Before the Cold War began, the Bolshevik Revolution in 1917 resulted in the overthrowing of the Tsars who had ruled Russia for years and relinquished power from the Tsars to the Bolsheviks, a socialist party (Cold War). The revolution and the ensuing events impacted artists of all disciplines greatly as it was mainly the social hierarchy and funding from social elites and royals such as Tsars that made it possible for artists to make a living. With the rise of the Bolshevik's socialist party, this social standing that made it possible for artists to support themselves would soon disappear. Notably, between the Bolshevik Revolution in 1917 and the second World War in 1939, roughly seven hundred and fifty artists left Russia and settled abroad (Вздорнов). This shift in social standing was one of the reasons both Balanchine and Stravinsky left Russia and became expatriates, never returning to their homeland once they left (Walsh) (Bentley, 101). As this paper will elaborate on, the collaboration of the two expatriates would result in a ballet that helped change the landscape of neoclassical

dance. The tensions between the U.S. and Russia during the Cold War period were very present during the creation of *Agon* in the late 1950s, especially because of both Balanchine and Stravinsky's connections to both the U.S. and Russia as a found home versus a birth home that they had both abandoned. Because of these complicated connections, the Cold War and U.S. Russia tensions can be seen as influences on the choreography of *Agon,* as the paper will later elaborate on.

The push and pull of *Agon*, which translates into "contest, struggle, or competition" from Greek, reflects the tensions between the U.S. and Russia at the time of its creation and can also be seen as a metaphor for Balanchine's love for his new home, America. This inherent drama of the choreography uniquely combined Balanchine's adherence to Russian balletic traditions with the cultural influences of modernism. Contrary to one of the only other major American ballet schools at the time, the American Ballet Theater School, Balanchine imported much of his staff for his school, the School for American Ballet, from Russia. It was a point of great difference between Balanchine's school, the School for American Ballet, and the other great American ballet school, the American Ballet Theater School, that the teachers Balanchine brought in were from distinctly non-American backgrounds. Balanchine's teachers had training in the Danish, English, and Russian classical traditions of ballet, while the American Ballet Theater's school was a uniquely American school with little to no outside influence whatsoever in their ballet training ("History"). Notably, Balanchine's costume designer and many pianists working at the School for American Ballet were Russian, which a former student of Balanchine noted made the school feel like a small Russian community in itself (Bentley). Balanchine also imported Russian dancers in keeping with the Russian standards of ballet. Ballet stars such as Baryshnikov came all the way from Russia to dance for Balanchine's new company, which added the the atmosphere that set the School for American Ballet apart; Balanchine aimed to create something Russian in his new home of America with all the standards of Russian ballet and training.

At the time of *Agon*'s creation and presentation to the general public in 1957, segregation was still in place in the United States. The Civil Rights Act that officially ended segregation across America was not put into effect until 1964, well after *Agon*'s premiere (Eldridge). The original casting of *Agon* was de-segregated, with the cast of the *pas de deux* consisting of Diana Adams and Arthur Mitchell– a white woman and an African American man (The Kennedy Center). This was unheard of for the time, since racial integration was still a tender topic to many Americans. In 1957, desegregation had occurred in the military which was a large step in the right direction towards desegregation as a nation, but the general public had still yet to embrace the notion, making Balanchine's casting progressive for the time (Dudziak, 3). The public's reaction to the interracial casting did not amount to much, as seen in a New York Times review of *Agon* published a month after *Agon*'s premiere, which failed to mention the fact that one of the principal soloists, Arthur Mitchell, was a person of color. After commenting that *Agon* was "danced, and danced magnificently," the writer concluded the review with the simple observation that "The chief soloists are Melissa Hayden, Diana Adams, Todd Bolender, and Arthur Mitchell, and they all deserve some sort of Congressional medal." (New York Times, 56) The lack of

reference to Arthur Mitchell's race gives the impression that the press and the public didn't care so much about the integrated casting rather than the dancing itself, which the writer suggested was good enough to win a Congressional medal, including Arthur Mitchell in this observation. This review leans positive in terms of the progressive casting because if the public was scandalized to see a dancer of color in the ballet, the review most likely would have included that detail or even language to suggest that a dancer of color in such a starring role was something that the public disapproved of. The integrated casting was a large step in the progressive direction. Many dancers and choreographers of color had fought for integration in dance before *Agon*'s creation. One of such dancers and choreographers was Katherine Dunham, a well known modern dancer and founder of the Katherine Dunham Company, who refused to perform at theaters that did not have integrated seating (Buis). Her fight helped bring dance to a larger audience, which paved the way for Balanchine's integrated casting because of the racial inclusivity it pushed in the dance world. Dunham's fight for inclusion also pioneered the struggle for recognition of non-white dancers, a struggle which continues today.

The shift in *Agon* from a strictly balletic style to a style more impacted by modern dance heralded the change in style of Balanchine's later works and some uniquely "Balanchine" moves. Modern dance had been growing in influence in the United States since the 1920s when Martha Graham first began developing the Graham technique– a style of modern dance that is one of the most influential styles to date and centers around the use of the pelvis and arms to create provocative lines. Many modern techniques are present in *Agon* and Balanchine carried them with him while creating other compositions throughout his career, such as the use of flexed hands and straight arms in leaps (Figure 1), kicks in parallel positions that reject the classical turned out ballet style, and throwing the torso out of alignment with the hips. As visible in a still shot from the 1957 recording of *Agon* performed by New York City Ballet (Figure 2), these moves contrast the traditional balletic standards of alignment. These innovations as displayed in *Agon* form a fundamentally unballetic style that takes the body out of alignment, the shoulders and pelvis not stacked on top of the hips as classical ballet demands. This move deviated greatly from Balanchine previous works, all of which were choreographed in classical Russian ballet style (Wilkin, 705). Such unballetic moves are seen repeatedly in later Balanchine pieces such as *Rubies* and *Slaughter on Tenth Avenue*, which were also choreographed in the neoclassical style that *Agon* heralded. The manipulation of the female body by a male figure is also seen in the works of Martha Graham (Figure 3), and was a concept that starkly contrasted the use of very gentle partnering in classical ballets. Balanchine's use of manipulation of the female body in *Agon* greatly contrasts the work of previous classical ballets and furthers the influence of modern technique on Balanchine's neoclassical works, specifically *Agon.*

Ultimately, the use of American modern techniques in *Agon* shows how Balanchine embraced the country that took him in after his exile from Russia. By blending the classical Russian ballet techniques with the modern techniques that America had both birthed and popularized such as the Graham technique and the Horton technique, Balanchine created *Agon,* a neoclassical masterpiece that helped change the landscape of neoclassical ballet. By embracing

the typically American styles such as Graham and Horton, Balanchine adapted the style that he was trained in, Russian ballet, to appeal to American audiences to create a style that was both stylistically and visually unlike anything anyone in the ballet world had seen before. His balletic innovation cemented him a place in American history not as a foreigner who made works in America, but more so as an innovator who created uniquely American ballets by embracing American styles of dance and linking them with his own. Through this distinct combination of influences and factors, Balanchine became known for his neoclassical style throughout his career with other ballets such as *Rubies (1967)* and *Stravinsky Violin Concerto (1972)*. In *Rubies,* Balanchine uses the motif of flexed hands and feet as well as manipulation of the female body (Figure 4), both of which motifs originate from *Agon*. *Stravinsky Violin Concerto* also carries the motif of flexed hands and feet, following *Agon's* neoclassical style of choreography as well as its unconventional and slightly unclassical style of music. Balanchine's style of modern dance and classical ballet also impacted the future of ballet, especially at Balanchine's company, the New York City Ballet. *Agon* helped pave the way for more neoclassical choreographers in ballet such as Kyle Abraham, Gianna Reisen and others who are included in the New York City Ballet repertory among the neoclassical works of Balanchine. New York City Ballet is arguably best known for the neoclassical works it has commissioned over the years after Balanchine's death in 1983, and a large portion of their programming is dedicated to works of neoclassical and contemporary ballet. *Agon* plays a part in this as the jumping off point for Balanchine's signature style and an inspiration for generations of neoclassical ballet to come. Balanchine's neoclassical works are still in circulation in the program of the New York City Ballet, which shows the timelessness and importance they give as some of the company's signature works. While there is no definitive number of times *Agon* has been performed, it is performed by ballet companies around the world such as the Paris Opera Ballet, Royal Swedish Ballet, and Norwegian national ballet to name a few (The George Balanchine Foundation). This serves as a reminder of Balanchine's influence as a choreographer worldwide, even after his death, as well as an example of the far reach of the neoclassical ballet movement that he pioneered.

The composition of *Agon* is arguably the most unique aspect of the piece. Balanchine choreographed *Agon* to a tailor-made piece of music composed by Stravinsky, and the majority of the movements in the ballet are tied directly to the emphasis and accents in the music Stravinsky created. This circumstance makes it virtually impossible to perform the choreography of *Agon* with any other piece of music. Balanchine's choreography is just as avant-garde as Stravinsky's music; the ballet does not appear to be choreographed in a clear measure of eight like almost every other ballet composition. Instead, it is strictly bound to the accents and cues that Stravinsky composed. Stravinsky's music is as unusual as the ballet itself. The composition made for *Agon* features many unclassical instruments such as the timpani (a large drum), castanets, and a xylophone, all of which combine to create the unclassical sound of the score (Keller, 1). By including such instruments, the music of *Agon* breaks classical tradition and conventionality just as the ballet does. In order to tie these unconventional instruments to the more commonly known classical instruments such a string instruments and piano, Stravinsky

notably aligned the changing rhythms of the odd instruments such as the woodwinds, harp, and strings with the constant, underlying dance rhythm that the ballet revolves around (Richardson, 327). This meshes the classical and unclassical sides of the ballet while making sure the music is a cohesive combination of both, and gives the music the sound of controlled chaos. This layers the ballet with tension before the dancing even begins as the instruments seem like they're fighting to be heard over each other, which aligns with the conflict and tension that the name of the ballet references. The relationship between the music and the dancers in *Agon* is one of a kind as well. Because the music is not composed in recognizable counts of eight, the dancers are controlled by the accents of the music– something almost never seen in classical ballet before *Agon*, though considerably more seen today. The unclassical music adds to the modern influences behind *Agon*, as before its creation it was mainly modern dancers who used music that was considered to be "unclassical." Balanchine and Stravinsky broke this stereotype when Balanchine created *Agon* as a neoclassical ballet using Stravinsky's unconventional music.

The costuming of *Agon* is also unique for its time. Both the male and female dancers wear very simple black and white costumes; black leotards and pink tights for the women and white shirts with black pants for the men. The simplicity of these costumes allows the audience to focus solely on the choreography of the ballet and the lines of the dancers without adding a layer of distraction that more complex or colorful costumes might pose. Simple costuming is not the norm for classical ballet performances, with most story ballets such as *The Nutcracker, Swan Lake,* and *The Sleeping Beauty* flaunting elaborate costuming including long skirts, tutus and elaborate multicolored leotards for the women and colored vests or elaborate cloaks for the men. Rather than resembling the costuming of classical ballets, *Agon's* costuming is more reminiscent of modern dance, where the costuming is kept to a minimum in order to showcase the lines of the dancers. This is another way in which Balancine embraces American traditions, using American styling that originated in modern dance instead of the elaborate classical ballet styling used in his home country of Russia. Balanchine pays homage to modern dance in more than the choreography of the ballet by borrowing the styling of modern dance as well. By leaning into the modern aesthetic while still using the basic ballet attire of a leotard and tights, the costuming advances the neoclassicism of the ballet in an aesthetic sense in addition to the technical and musical ways in which *Agon* is considered neoclassical.

The *pas de deux*, or dance of two, in *Agon* is the centerpiece of the ballet that reveals the most about Balanchine's mindset while he was choreographing. Although the ballet has no storyline, there is a distinct push and pull dynamic between the male and female dancers onstage that can reflect Balanchine's outside inspirations for the piece. The push and pull also alludes to the name of the ballet itself; *Agon* in Greek translates to a struggle or contest that is evident between the two dancers. At the time Balanchine choreographed *Agon* in the late 1950s, the tension between the U.S., Balanchine's adoptive home, and Russia, Balanchine's birthplace, was high (Schlesinger, 30). Politically, the United States and the Soviet Union perceived each other as rivals, as for each country the other was the only power in the system that posed a serious military threat to its own security. (Snyder and Diesing, 420) The push and pull dynamic in the

pas de deux between the male and female dancers, as well as the palpable asymmetry and unclassical influences, can be seen as a metaphor for the tensions between the two countries. The push and pull in *Agon* also draws reference to the way that conflict defines our lives, usually in multilayered ways. The surface push and pull that the movement provides in union with the conflict between Balanchine's Russian upbringing and American found home creates layers of tension in the ballet. The lack of symmetry in the pas de deux is also notable compared to the rest of the ballet, with the other sections always having even formations in which the number of dancers on each side of the stage is near balanced. In the pas de deux, the push and pull dynamic between the two dancers takes them all over the stage to form both classical and unclassical movements. This is comparable to well-known classical pas de deuxs from the Nutcracker's Sugarplum pas de deux and the Sleeping Beauty wedding pas de deux to every other pas de deux in classical story ballets, which all adhere to a strict theme of symmetry. Symmetry has been a longstanding theme of classical ballet. In fact, Russian theater critic Akim Volynsky wrote that "In order to be a balletmaster, it is necessary to be able to think in the geometric likenesses of movements and poses that serve as the core of all material compositions." (Volynsky, 20) This reinforces the importance of symmetry to the aesthetics of ballet. The symmetry in more traditional ballets is created between the two dancers and the circular motion of the *pas de deuxs*, where the two dancers meet upstage center and reunite to bring the pas de deux back to balanced form between the man and woman onstage. This only occurs at the very beginning of the pas de deux in *Agon*, where the two dancers meet in the middle of the stage and perform some of the only classical based steps in the whole pas de deux. The rest of the pas de deux is filled with push and pull motions that– rather than providing an equal dynamic between the two dancers– leads to a constant grapple for power, which results in a lack of symmetry as each dancer seems to move solely with the intent of gaining the upper hand. The pas de deux also includes a lot of male manipulation of the female body, which is seen in one of Balanchine's later works, Rubies (Figure 4). Balanchine's use of manipulation of the female body in *Rubies* is mainly one woman being manipulated by multiple men, who move one leg while holding the rest of her body still. In contrast, in *Agon* only one man is manipulating the body of one woman. Manipulation of the female body in *Agon* serves to further the dynamics of push and pull between the two dancers, as when the male dancer is manipulating the female dancer's body, she is ceding control of the pas de deux to him, shifting the power dynamics in favor of him. When she gains back control, she also is free to gain back the upper hand, and so the struggle for power that *Agon* is named after continues.

There is also a feminist lens that can be used to look at Agon. Broadly, Agon was choreographed by a man, and therefore it can be assumed that it was choreographed under the "male gaze." The "male gaze" is an idea that originated in feminist film theory and dictates that women when directed by men are displayed as more passive, and more conforming to the heterosexual male ideal of a woman (Daly 2). This concept is seen mainly in the pas de deux of Agon, where the pairing is made up of a man and a woman, thus satisfying the heterosexual portion of the male gaze. The manipulation of the female body by the man in the pas de deux is

also arguably a result of the male gaze because of the way it presents the woman as a passive subject to the man's movements and physical commands. The male gaze can also be seen in the pas de trois containing two men and one woman. The casting of the pas de trios (two men and one woman) is such that masculinity overpowers femininity two to one. The woman is never unaccompanied by the two men, and they serve as both onlookers and looming figures during her solo. In the last part of the pas de trois, male manipulation of the female body is seen again, this time with the woman having no control over her body. The men pick her up and throw her between them once in the beginning of the section and once at the end of the section. In both instances, the woman is stiff and compliant to the actions of the man. This can be interpreted as the male gaze because of the passiveness it gives the woman in the pas de trois, almost making her seem like a physical object rather than a human being. The men outnumbering her also undermine her presence by treating her as such, completing the ideal of a passive woman made for the male gaze. In addition to these moments in the ballet, Balanchine was also known to think of women as his "muses," often creating whole ballets for specific dancers whom he felt artistically drawn to (Bentley). Balanchine's expectations of a female dancer also exemplifies the reach of the male gaze in his ballets; the only women Balanchine put on stage met the high standards of being incredibly thin and flexible, standards which redefined what it meant to be a dancer in the 20th and 21st century due to Balanchine's influence over the dance world. All of the female dancers in *Agon* meet these flexibility and body standards, furthering the hold that the male gaze has over the ballet by adhering to Balanchine's standards that he set specifically for women. Zooming out, ballet can also be interpreted as an art form made specifically under the male gaze. One modern author describes ballet companies as "cults in which women are starved and controlled by men" (Kourlas). This phrasing backs up the notion that women in ballet have a lack of agency and control over their own bodies because of the larger role that men play in ballet companies. The author also mentioned the sexual abuse scandals and accusations that have plagued men in high ranking positions in various ballet companies, especially Balanchine's own company of New York City Ballet, which serve as reminders of the constant physical and sexual threats to women in art forms and sports such as ballet that value the silence and pliability of women before values such as self advocacy or honesty. The power that men have over almost every aspect of women in ballet companies reinforces the idea that ballet as an art form is generally dictated by the male gaze, as women are controlled in everything from how much they eat to what they do with their bodies. Another notable point when looking at ballet as an art form through the feminist lens is the notion that women are often objectified in the ballet world. The objectification of women can be defined as looking at women as objects or commodities rather than human beings with feelings. In the ballet world, it is commonly accepted that women are used for their body parts and physical attributes such as thinness or flexibility rather than any other qualities that they might possess. This view of women can be interpreted as the male gaze lending itself to the art form of ballet by recognizing that women don't have much agency in the way that their bodies are used in ballet, usually as marketing objects to display the larger art form of ballet in a way that is "aesthetically pleasing" to the men surrounding them. This

objectification can also tie back to the strict body standards that Balanchine held for women in his company and the expectation that he started that women be overly thin in order to be a successful ballerina in his company, an expectation which spread to the rest of the ballet world and persists today.

Due to both the collaboration between Balanchine and Stravinsky and the specific historical context, *Agon* was an extraordinary ballet. *Agon* helped to inaugurate a new genre of dance that Balanchine became known for: neo-classical ballet, which incorporated elements of un-balletic styles of dance such as jazz, contemporary, and modern. This style is seen in many of Balanchine's later works such as *Rubies*, *Slaughter on Tenth Avenue*, *Stars and Stripes*, and *Who Cares?* among others, and continues to be seen in the works of many contemporary ballet choreographers to this date. Not only does the ballet reflect the notion of push and pull physically, but it also reflects the push and pull tensions of the geopolitical state that *Agon* was made in, specifically the Cold War that the U.S. and Russia were enmeshed in. In addition to a political lens, *Agon* can also be examined through a feminist lens to show how the "male gaze" shaped key moments in the ballet such as the manipulation of the female body by a male figure. Balanchine's high standards for ballet also changed the ballet scene worldwide as many other ballet schools shifted to accommodate what Balanchine considered to be the "perfect dancer": Someone overly thin, lithe, and flexible, while exhibiting obvious technical perfection. These unrealistic body standards still influence the dance world today. Despite recent attempts to change these standards, the pursuit of the 'perfect Balanchine body' leads many dancers to disordered eating and decidedly unhealthy relationships with their bodies. The dance world today is left to deal with the unhealthy and overall negative effects of Balanchine's pursuit of perfection despite his choreographic legacy as one of the most inspired neoclassical choreographers of his time.

Figures

Figure 1: Still shot from George Balanchine's *Agon*

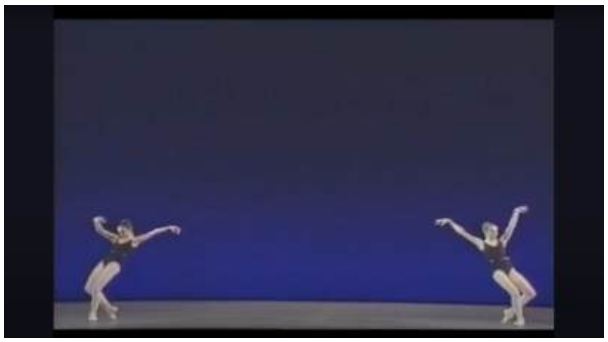


Figure 2: Still shot from George Balanchine's *Agon*

Figure 3: Still shot from Martha Graham's *Errand Into the Maze*

Figure 4: Still shot from George Balanchine's *Rubies*

Works Cited

A. Volynsky. Kniga likovanij. Prague, 1925, p. 20.

Bentley, Toni. Serenade: A Balanchine Story. Vintage, 2024.

Buis, Johann. “Dancing Against Segregation.” Wheaton Magazine, Winter 2021, Wheaton College, https://www.wheaton.edu/magazine/winter-2021/dancing-against-segregation/.

Daly, Ann. “Unlimited Partnership: Dance and Feminist Analysis.” Dance Research Journal, vol. 23, no. 1, 1991, pp. 2–5. JSTOR, https://doi.org/10.2307/1478691. Accessed 5 May 2026.

The Editors of Encyclopaedia Britannica. “Cold War.” Encyclopaedia Britannica, 27 Oct. 2025, https://www.britannica.com/event/Cold-War. Accessed 15 Nov. 2025.

Eldridge, Alison. “Civil Rights Act of 1964.” Encyclopaedia Britannica, 27 June 2025, https://www.britannica.com/story/civil-rights-act-of-1964-united-states. Accessed 15 Nov. 2025.

Feeney, Anne. Balanchine and Stravinsky: A Selected Bibliography and Catalog of Works. New York Public Library for the Performing Arts, 1982.

Garis, Robert. “Balanchine Institutionalized.” Salmagundi, nos. 50–51, 1980, pp. 234–41. JSTOR, http://www.jstor.org/stable/40547423. Accessed 30 Oct. 2025.

The George Balanchine Foundation. “316. Agon 1957.” Balanchine Catalogue, balanchine.org, https://balanchine.org/balanchine-catalogue/catalogue-of-works/316-agon-1957/. Accessed 15 Nov. 2025.

“George Balanchine + Agon.” The Kennedy Center, https://www.kennedy-center.org/education/resources-for-educators/classroom-resources/media-and-interactives/media/dance/george-balanchine--agon/. Accessed 15 Nov. 2025.

“History.” School of American Ballet, https://sab.org/history/. Accessed 23 May 2026.

Kourlas, Gia. “Can Balanchine Be a Feminist?” The New York Times, 5 Apr. 2023, https://www.nytimes.com/2023/04/05/arts/dance/balanchine-feminism-ballet.html.

Richardson, Mark. “A Bow to the Past: Seventeenth-Century Dance Rhythms in Stravinsky’s Ballet ‘Agon.’” The Musical Quarterly, vol. 97, no. 2, 2014, pp. 309–51. JSTOR, http://www.jstor.org/stable/43865513. Accessed 20 Nov. 2025.

Snyder, G. H., and P. Diesing. Conflict among Nations. Princeton UP, 1988, p. 420.

Walsh, Stephen. Stravinsky: The Second Exile (France and America, 1934–1971). Alfred A. Knopf, 2006.

Wilkin, Karen. “Remembering Balanchine.” The Hudson Review, vol. 46, no. 4, 1994, pp. 705–12. JSTOR, https://doi.org/10.2307/3852134. Accessed 16 Nov. 2025.

Вздорнов, Г. И. “Предисловие.” Художественная культура русского зарубежья, 1917–1939: Сборник статей, ВикиЧтение, 2005, https://culture.wikireading.ru/72480.

The New York Times, 28 Nov. 1957, p. 56. The New York Times TimesMachine, https://timesmachine.nytimes.com/timesmachine/1957/11/28/140147402.html?pageNumber=56.

The Untold Heroes Behind the Line in the Second World War By Daniel Meng

Abstract

Throughout the Second World War, many different armies and factors played a role in the fall of Nazi Germany. One lesser acknowledged aspect is resistance, especially in Western Europe. Though resistance impacted the war, the majority of people remained passive. However, the small actions committed by these resisters added up and saved a portion of the Jewish population and helped the Allied forces. This paper covers three different types of resistance: the rescue of Jews, passive resistance, and violent forms of resistance against Nazi Germany, focusing mainly in Western Europe. The multiple attempts to rescue Jews mainly happened in Denmark; still, other countries also contributed. Passive resistance included resistance through writing, smuggling, and other forms. Lastly, violent resistance included stealing, sabotaging, and destroying. Due to the cumulative efforts of resistance groups, the whole resistance in Europe forced Nazi Germany to fight a war on multiple fronts and from within, ultimately wearing down Germany.

Introduction

During a time when humanity seemed at its worst, for a brave few, complacency was not an option and resistance became a moral necessity. Throughout the Second World War, Nazi Germany committed atrocities against the Jewish population. However, a distinct minority in occupied countries rose and resisted the Axis occupation. Resistance during the Second World War played a vital role in undermining the Axis powers. Resistance took on many forms, including fighting against anti-semitism and protecting the Jewish population. Mainly through evacuation and hiding, thousands of Jews were saved by the collective efforts of ordinary citizens. Passive resistance was also a way of resisting Nazi propaganda. Through different publications and movements, citizens spread narratives that contested the false narratives of the Nazis. Lastly, many resistors were violent to the German occupiers, causing significant damage to their machinery and plans. Nevertheless, there were many other forms of resistance not covered in the paper.

In 1938, Hitler demanded the Sudetenland, part of Czechoslovakia with many ethnic Germans. At the Munich Conference, Neville Chamberlain, Prime Minister of the United Kingdom, gave in on the condition that he would not invade the rest of Europe. Despite his promise, Hitler went against the agreement at the Munich Conference and invaded the rest of Czechoslovakia. After Hitler signed the Nazi-Soviet pact, Hitler and Stalin invaded and occupied Poland. Hitler quickly invaded and soon occupied Denmark and Norway. In April 1940, the Nazis dropped propaganda leaflets to the public, and the next day, Christian X, the King of Denmark, surrendered without a single shot fired, and France fell later (Hoose 10). Hitler and the Nazis occupied many countries such as Greece, Yugoslavia, and Denmark. Within these occupied countries, millions of Jewish civilians and others were persecuted and forced into concentration camps, where they were worked to death.

Many credit the Allied soldiers with winning the war. Yet, people behind the lines played a critical role as well. While many historians recognize the resistance in Eastern Europe in the Soviet Union and Yugoslavia, Western Europe had a vital force of its own as well. Churchill nicknamed Denmark "Hitler's singing canary" in reference to the lack of a fight (Lampe). However, unbeknownst to him was an army of civilians fighting back behind the lines and causing damage. A minority of Europeans fought their occupiers. Every act of sabotage contributed to the Nazis' downfall.

Throughout history, sabotage has occurred in most wars. It often occurred in active and passive forms. Passive forms of resistance often included literature, propaganda, rallying the people, and other actions. These actions targeted the general public to sway their views. Similarly, in the Second World War, insurgents utilized this to fend off Nazi propaganda. This type of literature included information on the Jewish people and their achievements, nationalistic ideas, and battle news that the Nazis often restricted and censored. However, a form of new passive resistance had sprouted during this global conflict in response to the Holocaust and Hitler's merciless slaughter of the Jewish population. Thousands of people came together to save their neighbors and their countrymen, either by hiding them in their own houses, renting apartments for them, or by illegally smuggling them to neutral countries. Yet, during wartime, passive resistance was not sufficient to win a war behind the lines. Violence was necessary.

The Rescue of the Jews

During the time of the Nazi occupation of European countries, both individuals and large groups participated in activities that sought to save the Jewish populations within their home countries, as the citizens treated them as equals.

One of the biggest contributors to the movement to save the Jewish population was the creation of fake IDs and false personas. Women such as Maria von Maltzan of Germany made fake IDs for hundreds of Jews (Atwood). Additionally, a woman named Irena Sendler took in Jewish children and gave them Christian identities (Atwood). By doing this, she was able to save them by removing their Jewish identities. These two women directly saved hundreds of Jewish people while putting their own lives at risk. Henri Scharrer, another hero, created false identities for a variety of people, including Jews, English and American soldiers, and other people to escape or return to their countries (Scharrer). These efforts were employed by the few people who decided to assist in the evacuation of Jews. In addition, many people assisted the Jewish population in fleeing occupied areas into neutral countries such as Sweden.

Many Jews sought refuge in countries that were not occupied by Germany. However, this could not happen without the help of others. For example, Joop Westwill smuggled hundreds of Jews to Spain. A Jewish girl that he saved described him saying, "I'll never forget him. In those dark days, he was the only spark of humanity" (Stadtler 142). Even though he only saved a marginal part of the Jewish population, each life saved added up, resulting in tens of thousands of Jewish civilians escaping a horrible fate. Another example of this is Aristides De Sousa Mendes, a Portuguese diplomat who disobeyed orders to write thousands of visas for the Jewish

to get to Portugal. He later said about his actions, “I had to save these people-as many as I could. If I were disobeying orders, I would rather be with God against men than with men against God” (Stadtler 145). Because of his actions, Mendes saved thousands of Jews alone. Clearly, one of the biggest motivators to save the Jews was its moral necessity. People believed it was right to intervene. This sense of righteousness is what kept thousands of Jews alive.

Additionally, many groups also attempted to smuggle Jews out of occupied areas. For example, the Elsinore Sewing Club developed into a club that smuggled Jewish people out of Denmark. They believed that the Jews were not just Jews but also Danish (Slot 144). In other words, they were the same as everybody else living there, and they were not subordinates but rather closer to family, not believing in the Nazi propaganda that somehow the Jewish were subordinates. Due to the concerted efforts of these clubs and others, almost all the Jews living in Denmark were rescued. In addition to escape, many families and individuals decided to hide Jews in their homes, using up their own resources to take care of them.

Thousands of Jews were saved through these heroic actions, mainly in Denmark. Across Europe, civilians risked everything to house Jews too. An example of this can be found in Stefania Podgorska. She rented apartments under her own name for Jews, as she claimed that their lives were as important as hers (Atwood). She shows the importance of saving the Jewish population to the locals. As a result, she saved many lives. Many others followed, displaying the moral necessity of resistance, even at the expense of their own lives. Her mindset depicted that of many others who hid Jewish communities within their houses. As a result, many risked their lives to save those in danger.

In addition to these individuals helping, many groups, such as hospitals, helped as well. Danish ambulance drivers circled Jewish names in a phonebook and drove around to alert them, with many drivers hiding those who did not know where to flee in their own hospitals (Slot 144). Most of these drivers did it themselves, without any organized action, showing how they did it out of concern for the Jews and genuine love. Many doctors in hospitals, such as Karl Koster, hid Jews in their hospitals (Silver 3). His actions resemble those of many others as well, and with others saved a large number of Jews from a guaranteed death. Many schools also refused requests to hand over Jews, with one saying that they “could not, for reasons of principle, comply with the request” (Focke & Michman 97). These schools not only stood up to the Nazis but also saved many children and their families from the Holocaust. During this awful period in history, these heroes became the bright spots and saved hundreds of thousands of lives.

Their cumulative efforts saved thousands of Jews. In Denmark specifically, only 470 of the almost 8,000 Jews were captured and sent to the German camps (United States Holocaust Memorial Museum). After the war, of the original nine million Jews, only 1,500,000 remained in Europe, excluding the Soviet Union (Shapira & Keynan). Due to the nature of war, how much resistance actually played a role in saving them may be debated, but it is certain that resistance had a non-negligible role in the rescue.

Passive Resistance

Patriotism and nationalism also served as important means through which countries were able to channel resistance to the authority of the Nazi Party. Many different groups implemented this as an attempt to stem the development of Nazi Germany.

Clandestine literature involved illegal literature produced during the war. During the occupation, Germany repressed literature to blind the occupied countries and hide the crimes of the Third Reich. Hundreds of organizations were set up to spread literature to educate the populations. These pieces of literature often inspired others to encourage resistance or spread news. A Danish high schooler by the name of Arne Sejr wrote a list of Ten Commandments for passively resisting Nazi occupation on the first day of German occupation, listing tasks such as working slowly for the Germans and defending every person persecuted by them, ending his pamphlet with: "Join The Fight For Denmark's Freedom." (Lampe). Later, Sejr started the Students' Enlightenment Service, which provided Denmark with news about the war and German atrocities. During these harsh occupations, lots of propaganda literature also rose. Many poems, such as the "Van der Lubbe," named after a communist worker who set the Reichstag on fire, cursed the Nazis. Excerpts of these poems include, "Until today, Germans haven't learned how to speak properly and just know how to yell and shout" (Dewulf 9). Poems like this sparked nationalistic fervor, alleging that Germans are less developed than other countries, such as Denmark, and became extremely popular among the people. These clandestine literary works pitted people against the Nazis. They brought attention to the horrors of the Nazis and sparked resistance, both violently and passively.

However, other than literature, many other forms of patriotism occurred, such as demonstrations by the people. Many resisters used different methods to spark nationalism in the people. Martyrdom became a popular way, and other influential figures gave powerful speeches. For example, Bernardus Ijzerdraat rallied the people, and his resistance group became nicknamed "The House of Orange," referencing the ruling family of the Dutch (Dewulf 19). In addition to this, at his execution, he represented the government as he wore orange pajamas on the day of his execution to show loyalty to the throne rather than to Germany (Dewulf 19). His ideas significantly impacted the public as he sparked nationalism among the people. Even after his execution, he became a martyr for his ideas and his resistance, causing more people to join the resistance against the Nazis. A Danish priest and playwright named Kaj Munk committed similar acts. When asked to stop resisting the Germans, he replied, "Never Consider! That is to succumb to deliberations, increasing irresolution, and finally cowardice" (Dewulf 84). Through his actions, not only did he stand up to the Nazis, but he also publicly inspired others to do the same. Lastly, the government turned into another form of resistance, specifically in Denmark. The government never gave up on its people throughout the occupation. For example, the King sold King's Badges for citizens to show their dedication to the King rather than to Hitler and the Nazi Party. In addition to the King's Badges, the government also built upon the ideas of the population. After surrendering to Nazi Germany, they later stated to the public, "By radio it was announced that the Denmark government rejected the German ultimatum and that they were at

war with Germany" (Korte). While this announcement seems insignificant, it displays the government finally standing up to Nazi Germany. Before this declaration, the government simply let Hitler occupy Denmark.

Resistance was not just a masculine affair; it also included women and even students. One example is Sophie Scholl, a German university student. With her brother, they published a publication called "The White Rose" at the Ludwig Maximilian University of Munich. Scholl often wrote pamphlets and distributed them among her classmates. For example, she wrote "We will not be silent. We are your bad conscience: The White Rose will not leave you in peace" (Atwood). This quotation displayed her persistence in standing up to Nazi propaganda with her own writing, eventually leading to her execution. Her persistence and her impact became apparent after her execution. The university was graffitied with the words "Scholl lives! You can break the body, but never the spirit" (Atwood). This embodied the spirit of resistance as these ideas spread through the general public, causing others to rise and fight against German propaganda. After her death, she became a martyr and inspiration to her classmates.

The effects of literature during wartime remain important. However, many argue that the limitations caused the literature to be futile for the most part. One such historian, Jane Chapman, argues that the Germans imposed many limitations that caused the resistance movements to be futile, such as imposing limitations on travel and the distribution of ink, duplicators, and paper (Chapman 7). As a result, resisters had limited resources to print and had trouble distributing their papers. However, her argument can only be perceived as true to some extent. While many smaller production sites had failed, successful ones joined with major production sites to produce literature instead. Others obtained jobs to smuggle ink and paper to production sites and smuggled publications and literature to specific groups of people. As a result, resistance groups produced clandestine literature throughout the entire war, and the German limitations on their processes had a limited effect on the industry due to the persistence of the resisters.

Violent Resistance

While most people resorted to passive resistance such as clandestine literature or saving Jews, others directly confronted the Nazi army through violent upheaval in many of the territories which the Germans occupied. Historians estimate that almost three thousand acts of sabotage towards factories and fifteen hundred acts of sabotage towards the Germans occurred throughout the war (Chapman 12).

The first type of violent resistance involved sabotage, specifically, making German possessions useless. Sabotage became the goal of many people and resistance clubs. The Churchill Club was formed when teenage high school students felt shocked that their Danish government had surrendered to Hitler even though their neighbors, the Belgians, fought back. They started with acts of sabotage, as can be seen from the writing of their leader, Knud Pederson. He claimed that the club started by knocking down German signs and turning them around to cause misdirection (Hoose 30). While this seems insignificant, he later found that the people around them had noticed, and it demonstrated that the Danes would not let the Germans

take over easily. His influence on the rest of the citizens was great, as in the future, Denmark would become a resisting country. Later, he further recalled that they pried the radiator screens off of German trucks during Denmark's occupation and sometimes lit them on fire by throwing a match into the fuel tank (Hoose 47). Later, they continued to sabotage and burn many Nazi and German possessions. Pederson recalled that they burned down a train full of German planes and blueprints, destroyed many German offices, and set factories on fire (Hoose 53). The actions of sabotage by this club had a nonnegligible impact on the Germans, showing that even high schoolers could hinder the plans of Nazis. Other groups cumulatively destroyed German equipment and supply lines, successfully contributing to the war efforts. Because of their many acts of sabotage, German factories were set back. Additionally, the Churchill Club consisted of a group of children. Many other resisters came together to work and fight for their country's freedom.

While these inexperienced teens worked as a group, many people worked independently. One man was Niels Skov, another Dane who came to detest the Nazi occupation. In his first operation, he threw a match into the tank of a German military truck. After this moment, he wrote, "There are more matches where this one came from" (Hopkinson 44). While he knew the risks of his actions and knew he crossed a boundary and broke the law, he realized that deterring the Germans was much more important, and he was willing to continue. He embodied to the highest degree the mindset and persistence of many resistance workers as they sought to hinder and destroy German possessions. Niels Skov and many other resistance fighters had started with one simple action, but soon, resistance had burst into flames (Hopkinson 54). These singular instances of resistance had slowly led to an increase in violent resistance, becoming a constant thorn in the side of Nazi Germany.

Another violent group of resisters was organized by professionals, unlike the Churchill Club. This involved the Holger Danske 2 group, also known as the HD2, named after the legend that a hero would awake the day Denmark was threatened. This sabotage group worked in close coordination with the British Special Operations Executive (SOE) to coordinate violent resistance against the Germans. In one of their earliest actions, the HD2 successfully infiltrated and destroyed a German factory called Varde Steelworks, which generated steel for the German military (Hopkinson 267). This action had a large impact as it hindered the production of steel, which was crucial to military efforts. In addition to this, the effects of these violent resisting groups could be seen on non-resisters as well. One of the chief engineers in charge of rebuilding the steel production site ordered the repair workers to work more slowly. This is a clear demonstration that the resistance workers had spread their ideas to the population as well, and people were willing to sacrifice to support the movement.

While sabotage became one way of violent resistance, others often stole from the Nazis as well. Many people involved in resistance resorted to stealing from the Nazis. One prime example is the Churchill Club. After initially working on sabotage, they believed that they needed weapons to eventually stand up to the Germans. The club collected knives, guns, bayonets, and other weaponry by snatching and stealing from the Germans (Hoose 125). After

this, the club made bigger efforts, such as stealing a machine gun from a barracks (Hoose 66). Even though these small attempts to take weapons by the Churchill Club seem insignificant, they actually caused damage to the Germans. When adding the cost of many other resistance groups, the amount of damage caused to Nazi Germany impacted the war. Not only this, but also these teenagers and the first acts of resistance had been extremely effective in inspiring the rest of the population, as eight times more resistance occurred in 1943 than in 1942 in Denmark (Hoose 143). Evidently, the Churchill Club influenced the rest of the population, causing damage to the German army. With the accumulation of multiple acts of sabotage, the overall damage done to the Germans was much greater than others believed. While this club played a key role in Denmark, other people had also contributed. The Knuckles Gang developed into a gang that stole food rations to feed the Jewish population (Slot 146). While insignificant to the Germans, it instead helped feed the hidden Jewish population that sought refuge during the Holocaust. Similar to the Knuckles Gang, many women contributed to taking these food rations during wartime as well (Atwood). Both of these acts helped the Jewish population, but minimally impacted the Germans.

The Churchill Club, the Knuckles Gang, and multiple women stole physical objects such as weapons and food rations. Another less-known type of resistance was intelligence operations aimed at taking information from the Germans and spreading it to the British and various other countries. Many people partook in intelligence operations, smuggling data and information on the British out of occupied areas. Tommy Sneum, a Danish flight officer, risked his life and took photos and videos of the new German Freya Radar system and repaired an old plane to sneak the intelligence report to Britain (Ryan). Seemingly not important, when combined with other intelligence reports from other spies, it provided crucial information to the Allied Powers. In addition to this, Sneum later conducted many other undercover intelligence operations throughout the war after escaping to Britain, supplying them with key information throughout the war (Holbraad 6). Due to the efforts of Sneum and numerous other agents employed by Britain and other agencies, major powers kept tabs on the plans of the Germans and their new pieces of technology. This espionage clearly gave the Allies an upper hand in their battles, knowing where the Germans would strike. Besides sabotage, many people and groups resorted to derailing and harming the Nazis by stealing. This method proved effective as they smuggled information out of occupied areas, armed themselves with weapons, and provided resources for the Jewish population that needed to be saved. However, even though sabotage and looting became common tactics among resistance groups, violence was still a necessity.

Even though resistance somewhat hindered the Nazis, some historians still doubt the extent of the effects. One historian, Tony Judt, argues that many parts of resistance became myths and were largely exaggerated after the war (Judt 6). Though it may be true that most people did not partake in the resistance movements, it is crucial to recognize that it played a key role in the war. For example, when the German police caught key members of the Churchill Club and found their stash of weapons, they estimated that it cost almost two million kroner in weapons and ammunition. This estimated amount cost the Germans almost five hundred

thousand dollars in the present day in the United States. While seemingly insignificant, this number only estimated the damage from a few untrained high school saboteurs and only takes into account stolen property, not the damage and destruction they had done to German military vehicles and equipment. Additionally, the damage done by multiple groups added up.

The previously mentioned Holger Danske 2 (HD2) soon turned into a military unit. Their military attacked members of the Danish guards and those serving the Germans, believing that they would be part of the military unit that liberated their home country (Holbraad 105). While being insignificant on a small scale, the collective efforts of multiple military groups significantly affected the effectiveness of the German army. As a result, the Allies greatly benefited as Hitler struggled to fight a war on multiple fronts against the rest of the world and from within.

Though many uprisings made an impact, Henri Michel argued that most resistance groups failed because they could not obtain Allied help (Michel). However, this statement is partially false, as many resistance groups did succeed in their goals. Throughout the war, groups such as the Holger Danske 2 and the Churchill Club had continuously irritated the Nazis through their actions. In addition, many uprisings admitted that they knew they would not work, but the point was not to win, but rather to fight and slow the Nazis down. Looking at the resistance from this perspective, most uprisings in fact succeeded. As a result, Michel's argument may be factually correct, but it does not take into account the actual goals of resistance movements during the war.

Conclusion

The Allied conventional forces are often credited for winning the war, with generals and leaders such as Churchill and Eisenhower credited with being war heroes. While this is true, resistance played an important role that must not be forgotten. A few citizens contributed significantly in this fight behind the lines with different types of resistance. First, in many occupied countries, ordinary citizens stepped up to protect Jewish people from the Holocaust. Many risked their own safety to shelter Jews, create forged identities, and smuggle them to neutral nations. Men and women joined this collective effort, with many risking their own lives and resources to save those around them. Even in humanity's darkest hours, acts of courage and compassion shone brightly. Amidst these heroic actions, the underground press emerged to fight Nazi propaganda. While Hitler introduced propaganda against the Jews and censored news to depict him as a hero, underground newspapers rose to combat this. People of all ages, from high schoolers to older generations, came together to inform the public of real information without Hitler's version. Lastly, resistors employed violence. The use of sabotage, robbery, and actual violence caused a lot of damage to Hitler and Nazi Germany. With the combination of all three factors, the impact of resistance clearly played an instrumental role in the outcome of the Second World War. Holistically, these acts of resistance eased the war for the conventional Allied forces to defeat the Axis. During the bleakest hours of history, resistance and defiance by ordinary people preserved humanity.

Works Cited

Atwood, Kathryn J. Women Heroes of World War II: 32 Stories of Espionage, Sabotage, Resistance, and Rescue. 2nd ed., Chicago Review Press, 2019. Google Books, books.google.com/books/edition/Women_Heroes_of_World_War_II/yN_17svfqSQC?hl=en&gbpv=0.

Chapman, Jane. Writing Resistance in the Second World War: Secrecy and Participation in Newspapers. Palgrave Macmillan, 2024. SpringerLink, https://doi.org/10.1007/978-3-031-60742-4.

Dewulf, Jeroen. Spirit of Resistance: Dutch Clandestine Literature During the Nazi Occupation. Camden House, 2010. JSTOR, www.jstor.org/stable/10.7722/j.ctt81vds.

Focke, J. W., and Dan Michman. "1940 to 1942: Occupation, Oppression, Persecution." Machseh Lajesoumim: A Jewish Orphanage in the City of Leiden, 1890–1943, Amsterdam University Press, 2021, pp. 95–134. JSTOR, https://doi.org/10.2307/j.ctv1zvccg1.12.

Holbraad, Carsten. Danish Reactions to German Occupation: History and Historiography. UCL Press, 2017. JSTOR, https://doi.org/10.2307/j.ctt1mtz53j.

Hoose, Phillip. The Boys Who Challenged Hitler. Farrar, Straus and Giroux, 2015. Google Books, books.google.com/books/edition/The_Boys_Who_Challenged_Hitler/62IGCAAAQBAJ?hl=en.

Hopkinson, Deborah. Courage and Defiance. Scholastic Press, 2015. Google Books, books.google.com/books/edition/Courage_Defiance_Stories_of_Spies_Sabote/SGlzBgAAQBAJ?hl=en.

Judt, Tony. Postwar: A History of Europe Since 1945. Penguin Publishing Group, 2006. Google Books, books.google.com/books/about/Postwar.html?id=vcqXXIifR40C.

Korte, Suzanne. Tragedy & Betrayal in the Dutch Resistance. Pen & Sword Books, 2020. Google Books, books.google.com/books?id=GBMZEAAAQBAJ.

Lampe, David. The Danish Resistance. Pickle Partners Publishing, 2016. Google Books, books.google.com/books?id=tIvjDAAAQBAJ.

Michel, Henri. The Shadow War: Resistance in Europe, 1939–1945. Translated by Richard Barry, Deutsch, 1972. Open Library, openlibrary.org/books/OL5761240M/The_shadow_war.

Ryan, Mike. The Hornet's Sting: The Amazing Untold Story of World War II Spy Thomas Sneum. Skyhorse Publishing, 2009. Google Books, books.google.com.pa/books?id=LkT-DQAAQBAJ.

Scharrer, Jos. The Dutch Resistance Revealed. Pen & Sword Books, 2018. Google Books, books.google.com/books/edition/The_Dutch_Resistance_Revealed/IrTNDwAAQBAJ.

Shapira, Anita, and Irit Keynan. "Survivors of the Holocaust: An Overview." Yad Vashem, www.yadvashem.org/articles/general/the-survivors-of-the-holocaust.html. Accessed 25 Aug. 2026.

Silver, Daniel B. Refuge in Hell: How Berlin's Jewish Hospital Outlasted the Nazis. Houghton Mifflin, 2003. Google Books, books.google.com/books/edition/Refuge_in_Hell/qqhIpd6CQ5YC?hl=en.

Slot, B. "The Resistance." The Annals of the American Academy of Political and Social Science, vol. 245, 1946, pp. 144–48. JSTOR, www.jstor.org/stable/1024814.

Stadtler, Bea. The Holocaust: A History of Courage and Resistance. Behrman House, 1994. Google Books, books.google.com/books/about/The_Holocaust.html?id=YvsD7WXNkMQC.

United States Holocaust Memorial Museum. "Rescue in Denmark." Holocaust Encyclopedia, 23 Sept. 2025, encyclopedia.ushmm.org/content/en/article/rescue-in-denmark.

Impact Of Family And Peers On Social Identity Formation In Adolescence

By Aanya Ghosh

Abstract

Adolescence is a crucial developmental stage in which people engage with their social surroundings to create a sense of self. Family ties, peer groupings, and educational environments all have a big impact on identity formation, which goes beyond personal psychological processes. Adolescents can explore their personal values, beliefs, social positions, and future goals in these social environments while forming a cohesive self-concept. While conflicting expectations in many contexts can lead to difficulties with identity negotiation and role conflict, positive social relationships support the formation of a healthy identity. Using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, this paper summarizes the body of research on the influence of peers and family on the development of social identities during adolescence. Academic resources such as PubMed, Google Scholar, ERIC, PsycINFO, SpringerLink, Taylor & Francis Online, and Wiley Online Library were used to find peer-reviewed papers. Social Identity Theory and Symbolic Interactionism, which offer theoretical viewpoints on how social interactions, group membership, and interpersonal relationships affect teenage identity formation, serve as the review's guiding principles. Adolescence and identity development, family and identity formation, peer group and adolescent identity, identity formation in school, and identity negotiation and role conflict are the five themes that comprise the findings. Peer groups enhance social belonging, self-esteem, and identity validation. Supportive family relationships foster identity exploration, autonomy, and emotional security. Through interactions between teachers and students, school belonging, and chances for social and academic engagement, educational environments further form identity. In order to integrate various expectations into a cohesive sense of self, adolescents also negotiate numerous identities in the contexts of their families, peers, and schools. Adolescent identity formation is a multifaceted, dynamic process that is influenced by ongoing encounters with important social contexts. Professionals working with teenagers, like teachers, parents, and psychologists can find value in knowledge about the effects discussed.

Keywords

Adolescent Identity, Social Identity, Family Influence, Peer Influence , Identity negotiation, Role Conflict, School Identity, Social Development, Social Identity Theory, Symbolic Interactionism, Identity Exploration and Commitment

1. Introduction

During adolescence, young individuals begin to explore their beliefs, values, aspirations, and future objectives as they transition from childhood dependence to more independence and responsibility. Family, friends, schools, and the greater social environment shape one's identity (Orenstein and Lewis, 2020). Before they have a strong sense of who they are, people try out a

variety of roles and philosophies. By describing identity development through the processes of commitment and inquiry, Marcia's Identity Status idea expands upon this idea and identifies four identity statuses: identity diffusion, identity attainment, moratorium, and foreclosure. These theories stress that as identity creation is a dynamic developmental process, teenagers should consider their options before making major personal commitments (Branje et al., 2021). Identity formation is dynamic and keeps changing as a result of interactions with shifting social contexts. Instead of happening as a single developmental event, identity is continuously shaped via self-reflection, dedication, investigation, and adaptability to new experiences. Identity formation is a lifelong developmental process (Branje et al., 2021). Family connections offer emotional support and direction and parents expose teenagers to cultural values, customs, beliefs, and behavioral expectations. Adolescents who have supportive familial relationships are encouraged to experiment and may have a stable emotional base. While restrictive family circumstances may impede identity development, open communication, parental affection, and opportunity for autonomy foster stronger identity commitment. Family relationships continue to shape teenagers' values, sense of self, and comprehension of social roles as they get older (Campbell et al., 1984). By offering academic, social, and extracurricular experiences that mold teenagers' self-concept and future goals, schools also serve as a crucial setting for identity development. Students' confidence, motivation, academic identity, and sense of belonging are all influenced by their daily interactions with teachers and peers. Adolescents are encouraged to develop competence, pursue their interests, and actively engage in school communities when there are positive teacher-student connections and supportive school environments. Students develop identities linked to success, leadership, and social responsibility as a result of these experiences (Roorda et al., 2011). Adolescence and identity development, family and identity formation, peer group and adolescent identity, identity formation in school, and identity negotiation and role conflict are the five main issues that are examined in this review, which is informed by Social Identity Theory and Symbolic Interactionism.

2. Methodology

2.1 Research Design

The study adopted a systematic literature review approach to examine the impact of family and peers on social identity formation during adolescence. The review used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. This was to ensure a transparent, structured, and replicable process for identifying, screening, and selecting relevant literature. Electronic databases used are PubMed, Google Scholar, ERIC, PsycINFO, SpringerLink, Taylor & Francis Online, and Wiley Online Library. The research used Peer-reviewed journal articles, review papers, and scholarly books published in English.

2.2 Inclusion Criteria

- Research on teenagers between the ages of 10 and 19.

- Focus on identity-related processes, social identity formation, or identity development.
- Examination of how identity formation is impacted by social relationships, schools, peers, and family.
- Use of scientific papers, systematic reviews, or peer-reviewed pieces released in English.

2.3 Exclusion Criteria:

- Research on targeted adults or children who are not in the adolescent age bracket.
- Discussion healthcare procedures that have nothing to do with identity formation.
- Non-peer-reviewed sources, editorials, dissertations, or conference abstracts.
- Examination of evolution of social identities or identity formation.

2.4 Study Choice

The four steps of the PRISMA process were inclusion, eligibility, screening, and identification. Titles and abstracts were used to filter articles found through database searches. Before full-text articles were evaluated for eligibility, duplicate records were eliminated. For the final review, studies that satisfied the inclusion criteria were chosen.

Each study's author, year of publication, research aims, methodology, sample characteristics, and significant findings pertaining to the development of adolescent identities were retrieved. A theme analysis technique was used to examine the chosen literature. The results were categorized into five main themes:

- Adolescence and Identity Development
- Family and Identity Formation
- Peer Group and Adolescent Identity
- Identity Formation in School
- Identity Negotiation and Role Conflict

How social interactions and group memberships contribute to identity formation during adolescence, the topics were analyzed via the lenses of Symbolic Interactionism and Social Identity Theory. This PRISMA-guided review offers a thorough summary of the body of research on how peers, family, and school settings influence teenage social identity.

3. Themes

1. Adolescence and Identity Development

This theme looks at adolescents as an important stage in development that is characterized by self-exploration and self-reflection. It also examines making a commitment

towards oneself. Theories used in this theme include: Marcia's Identity Status theory and Erikson's Identity versus Role Confusion (Branje et al., 2021).

2. Family and Identity Formation

The identity development of adolescents is affected by family processes, parenting styles, communication, values, and parent-child relationships. Families promote independence and create the foundation for values, security, and belonging (Benson and Johnson, 2009).

3. Peer Group and Adolescent Identity

This issue focuses on how adolescents' social identities, self-esteem, sense of belonging, and behaviors are influenced by their friendships and peer groups. Peer relationships offer chances to explore one's identity, but they can also put pressure on one to fit in (Dumas et al., 2026).

4. Identity Formation in School

This theme investigates how teenage identity is shaped by schools, instructors, peers, and academic experiences. Students' sense of competence, belonging, accomplishment, and future goals are influenced by their school settings (Roorda et al, 2011).

5. Identity negotiation and Role Conflict

This theme looks at how teenagers deal with conflicting demands from classmates, family, school, and society as well as numerous identities. It focuses on how young people create a cohesive sense of self by integrating various roles and experiences (Branje et al., 2021).

4. Discussion

4.1 Adolescence and Identity Development

During the crucial developmental stage of adolescence, people integrate their personal ideas, values, aspirations, and social duties to create a cohesive sense of identity. The Identity versus Role Confusion stage of adolescence, according to Erikson's theory of psychosocial development, is when young people actively investigate various roles, relationships, and life options in an effort to respond to the central query, "Who am I?" If this stage is completed successfully, a stable and cohesive identity is developed; if it is not, confusion, ambiguity, and trouble making judgments in the future may occur. Significant biological, cognitive, emotional, and social changes impact identity formation during adolescence. Adolescents with superior cognitive skills are better able to reflect on themselves, think abstractly, and assess their goals and views. They assess all facets of themselves and make well-informed decisions about their schooling, careers, relationships, and personal beliefs (Orenstein and Lewis, 2020).

Identity is a process of commitment and exploration that can change. Adolescent identity includes social identities like race, gender, culture, religion, and group memberships. Increased

psychological well-being, higher self-esteem, stronger self-efficacy, and better social and academic adjustment have all been linked to the development of a positive and integrated social identity. Adolescence is a critical time when identity formation establishes the groundwork for sound psychosocial growth and a smooth transition into adulthood (Orenstein and Lewis, 2020).

People start to actively examine their beliefs, values, ambitions, relationships, and future aspirations during this period, moving beyond childhood roles. Since young people try to acquire a cohesive sense of who they are and how they fit into their social settings, identity development is seen as one of the primary developmental goals of adolescence. Identity building is a process that requires both commitment and investigation. Before making decisions that support a stable sense of self, adolescents consider a variety of options including relationships, school, careers, cultural values, and personal views. According to research, identity formation is a dynamic process that is impacted by ongoing interactions between people and their social environments rather than a linear one. Peer relationships, educational experiences, cultural settings, and family ties are all significant sources of input that influence how teenagers perceive themselves (Branje et al., 2021). The process by which adolescents understand themselves and the way in which other people perceive them depends upon their different social interactions. This identity-making process is helpful for psychological well-being, self-respect, and a sense of belongingness. Adolescence is an important period of life where people construct and reconstruct their identities and develop the psychological tools needed for adulthood (Branje et al., 2021).

4.2 Family and Identity Formations

Since families are the initial social environments where people learn values, beliefs, cultural customs, and social expectations, they play a crucial role in the construction of adolescent identities. Young individuals start investigating their own identities during adolescence, including their aspirations, convictions, and potential responsibilities. However, interactions with parents and other family members, who have an impact on teenagers' perceptions of themselves and their role in society, affect this investigation. Healthy identity development is facilitated by supportive parent-adolescent connections. Positive family interactions, open communication, and parental support all promote identity development and commitment. Young people gain confidence and a better sense of self from families that offer emotional stability while giving them opportunities for independence. Adolescents' ties to culture, customs, language, and common values are shaped by their families, which can have an impact on social identity. Adolescents learn how to relate to wider social groups and gain a sense of belonging through family experiences. However, teenagers may negotiate and reinterpret parts of their family-based identities as they engage with classmates, schools, and larger communities (Crocetti et al., 2016). Identity development and self-reflection in female teenagers emphasized on how the capacity to reflect on one's own experiences affect the formation of a solid sense of self. Variations were considered among teenagers during several phases of identity development, such as identity attainment, moratorium, foreclosure, and diffusion, using Marcia's identity status framework. Teenagers who have more fully formed identities are better able to reflect on

themselves. They make meaningful decisions about who they are and who they want to become by using reflection to assess their beliefs, experiences, relationships, and future objectives. They combine many facets of themselves and strengthen their sense of self by thinking back on these events. Self-reflection as a crucial psychological process in the development of teenage identities, stressing that identity is formed via active investigation, assessment, and individual meaning-making (Campbell et al., 1984). A solid identity is developed in part through the quality of familial interactions, which include communication, emotional intimacy, parental support, and chances for independence. Adolescents who have supportive family relationships are encouraged to experiment with various ideologies, aspirations, and life decisions while still feeling secure. Adolescents form personal commitments and gain confidence in their ability to make decisions when parents respect their independence while offering direction. Teenagers can find it difficult to establish an independent sense of self in overly restrictive familial contexts. Family influence is closely linked to the daily interactions between teenagers and their parents. It goes beyond family structure. Positive communication, shared responsibility, and emotional support form an environment that encourages maturation and identity development (Benson and Johnson, 2009).

4.3 Peer Group and Adolescent Identity

During adolescence, peer interactions become more crucial as young people start to form identities in larger social contexts and strive for more independence from their families. Adolescents can explore many facets of themselves, get social feedback, and feel like they belong in peer groups. Thus, contacts with friends and social groupings, in addition to introspection, influence identity formation. Adolescents who have supportive connections can talk about their personal values, views, aspirations, and experiences while being accepted and encouraged by others. They gain confidence in experimenting with various roles and making commitments to their personal identities. By lowering feelings of loneliness and enhancing social connections, peer attachment, friendship quality, and involvement in peer groups support identity exploration. By influencing teenagers' perceptions of belonging and group membership, peer groups can have an impact on social identity. Teenagers frequently make comparisons between themselves and their classmates, forming parts of their identities based on common behaviors, attitudes, interests, and social experiences. The social ties foster beneficial growth by encouraging emotional health and acceptance. Peer pressure, however, can also result in pressure to fit in with the group's expectations, especially when acceptance starts to depend on adopting particular attitudes or behaviors (Ragelienė, 2016). Membership in social groups helps people form a portion of their self-concept, according to Social Identity Theory. Young people identify more with their peer groups during adolescence. These connections help them comprehend their own traits, ideals, and social standing. Tarrant investigated how teenagers assess their own peer groups in relation to groups to which they do not belong. The results show that teenagers have a tendency to form more favorable opinions of the groups they associate with, believing that their own peer groups possess more desirable traits. Greater affiliation and favorable opinions of a peer group are linked to stronger identification with that group (Tarrant, 2002). During

adolescence, peer groups have a greater impact on young people as they build their identities, become independent, and seek approval. Adolescents' reactions to peer pressure and control are influenced by their identity development. Two crucial stages are involved in identity development: identity exploration, in which teenagers consider various values, beliefs, and aspirations for the future, and identity commitment, in which they create a more distinct sense of who they are. Adolescents who have a strong sense of self can better handle the difficulties brought on by peer pressure. Teenagers who establish distinct personal values and commitments are more equipped to make choices that are consistent with their own convictions as opposed to merely conforming to the expectations of their peers. While identity exploration helps teenagers make decisions and avoid engaging in risky behaviors, identity commitment lessens exposure to harmful peer influences (Dumas et al., 2026).

4.4 Identity Formation in School

Adolescents continue to develop and negotiate their identities outside of the influence of family and close peer groups at schools, which are significant social settings. Students spend a lot of time in educational environments during adolescence. The interactions with teachers, classmates, and school communities help them understand who they are, what they can do, and what they can do in the future. Through extracurricular activities, social interactions, academic experiences, and role-playing, they can explore their identities better. Pupils gain confidence and a sense of competence with supportive teachers who offer encouragement, praise, and helpful criticism. By influencing students' perceptions of their skills and prospects for future success, positive interactions with professors can help students develop stronger academic identities. Teenagers' perceptions of their own skills may be impacted by unfavorable expectations or a lack of support, which could lower their level of learning engagement. Identity formation is directly linked to a strong sense of school belonging. Students form closer bonds with their peers and the larger school community when they feel valued, respected, and welcomed in their educational setting. Positive social identity development, emotional health, and self-esteem are all supported by this sense of belonging. Teenagers' perceptions of their social standing, competency, and belonging are influenced by these events (Verhoeven et al., 2019). A meta-analysis of 99 research with children ranging from preschool to high school revealed the connection between emotional teacher-student interactions and students' engagement and achievement. Positive teacher-student connections are linked to higher levels of school engagement and better academic results. Students are more likely to actively participate in their education and form positive self-perceptions when they get warmth, encouragement, and acceptance from their professors. By letting students know that their efforts and abilities are appreciated, motivating teachers can strengthen their sense of competence. Students regard themselves as competent learners and members of a school community. This becomes their academic identity. Negative teacher-student interactions can make children feel less like they belong and have a detrimental impact on their involvement in school. Teachers have a part in the social and emotional growth of teenagers in addition to providing academic teaching. Students' self-concept, drive, and

confidence are shaped by the feedback teachers provide them in their everyday encounters (Roorda et al, 2011).

4.5 Identity Negotiation and Role Conflict

Young people's cognitive development enables them to understand that they might act differently in various contexts. For instance, an adolescent can think of themselves as self-reliant and self-assured with friends, but more reliant or accountable around relatives. Teenagers may wonder which version of themselves best reflects their "true" identity as a result of these variances. According to Harter et al., because teenagers are learning to integrate many facets of their personalities and social expectations, confrontations are frequently more intense across various social roles. When teenagers encounter conflicting demands from various relationships, role conflict may arise. While friends may support diverse behaviors, interests, or social identities, family members may support particular ideals and obligations. Schools may set new standards for future objectives, leadership, and accomplishment. Adolescents must consider their own values and progressively create a more cohesive sense of self in order to manage these disparate expectations. When teenagers start approaching maturity, their capacity to blend numerous identities also increases. People progress toward coming to understand that several facets of themselves can coexist and contribute to a larger identity. It is not only seeing diverse roles as inconsistencies. Adolescents who have supportive interactions in a variety of social circumstances are better able to build their self-esteem, confidence, and sense of authenticity (HARTER et al., 1997). Exploration and commitment processes are part of the identity building process. Before making more firm commitments that are consistent with their personal values, adolescents investigate many options concerning school, careers, relationships, beliefs, and personal objectives. Through this process, young people can reevaluate their identities from childhood and create identities that more accurately reflect their growing independence and goals for the future. Adolescents get feedback, encouragement, and chances to consider many facets of themselves through relationships. Significant events and life changes can also have an impact on identity development by prompting teenagers to reevaluate their objectives and social positions (Branje et al., 2021). The Multiple Worlds Model, developed by Phelan, Davidson, and Yu (1998), describes how teenagers negotiate and manage the relationships, expectations, and values found in these various contexts. Identity formation is influenced differently by peers, family, and school. When the norms and values of each setting are the same, transitioning between various social realms can be easy. However, when teenagers encounter disparities in peer cultures, school requirements, and family values, identity problems may arise. Students try to strike a balance between peer relationships, academic objectives, and cultural customs or familial obligations. For teenagers from a variety of social and cultural backgrounds, identity negotiation is particularly important since they may have to deal with varying expectations in their homes, schools, and communities. Adolescents who successfully navigate these changes are able to maintain relationships with significant social groups while creating adaptable and cohesive identities (Davidson et al., 1998).

5. Conclusion

Adolescence and identity development, family and identity formation, peer group influence, identity formation in school, and identity negotiation and role conflict are the five themes examined in this review. These themes show how various social relationships interact to influence how teenagers perceive themselves and their role in society. The foundation for identity discovery is laid during adolescence, when young people start to form their own values, beliefs, aspirations for the future, and goals. Families offer emotional support, cultural values, and direction. This affects teenagers' early self-perception. Teenagers who grow up in supportive home circumstances are more confident and have a more solid identity because they are encouraged to be independent while yet feeling like they belong. Peer groups have a bigger role in forming social identities as teenagers strive for greater independence. Peer relationships and friendships have an impact on one's sense of self, sense of belonging, actions, and decisions. While negative social pressure may make it difficult for teenagers to uphold their personal ideals, positive peer interactions enable them to explore many facets of themselves. Schools also aid in identity formation of teenagers. Their academic and social identities are shaped by their relationships with professors and classmates (Verhoeven et al., 2019). Teenagers negotiate many responsibilities and expectations from classmates, family, schools, and personal objectives. All of it must work in integration. This can give adolescents the experiences and support they need to comprehend who they are and get ready for adulthood.

Works Cited

Branje, Susan, et al. “Dynamics of Identity Development in Adolescence: A Decade in Review.” Journal of Research on Adolescence, vol. 31, no. 4, Nov. 2021, pp. 908–27, https://doi.org/https://doi.org/10.1111/jora.12678

Benson, Janel E., and Monica Kirkpatrick Johnson. “Adolescent Family Context and Adult Identity Formation.” Journal of Family Issues, vol. 30, no. 9, Mar. 2009, pp. 1265–86, https://doi.org/10.1177/0192513x09332967

Campbell, E, et al. “Familial Correlates of Identity Formation in Late Adolescence: A Study of the Predictive Utility of Connectedness and Individuality in Family Relations.” Journal of Youth and Adolescence, vol. 13, no. 6, Dec. 1984, pp. 509–25, https://doi.org/10.1007/BF02088596

Crocetti, Elisabetta, et al. “Identity Processes and Parent-Child and Sibling Relationships in Adolescence: A Five-Wave Multi-Informant Longitudinal Study.” Child Development, vol. 88, no. 1, May 2016, pp. 210–28, https://doi.org/10.1111/cdev.12547

Davidson, Ann Locke, et al. Resources in Education. Teachers College Press, 1999, https://eric.ed.gov/?id=ed422060

Dumas, Tara M., et al. “Identity Development as a Buffer of Adolescent Risk Behaviors in the Context of Peer Group Pressure and Control.” Journal of Adolescence, vol. 35, no. 4, Aug. 2012, pp. 917–27, https://doi.org/10.1016/j.adolescence.2011.12.012

HARTER, SUSAN, et al. “The Development of Multiple Role-Related Selves During Adolescence.” Development and Psychopathology, vol. 9, no. 4, Dec. 1997, pp. 835–53, https://doi.org/10.1017/s0954579497001466

Orenstein, Gabriel A., and Lindsay Lewis. “Erikson’S Stages of Psychosocial Development.” PubMed, StatPearls Publishing, 7 Nov. 2020, https://pubmed.ncbi.nlm.nih.gov/32310556/

Ragelienė, Tija. “Links of Adolescents Identity Development and Relationship with Peers: A Systematic Literature Review.” Journal of the Canadian Academy of Child and Adolescent Psychiatry, vol. 25, no. 2, 2016, pp. 97–105, https://pmc.ncbi.nlm.nih.gov/articles/PMC4879949/

Roorda, Debora L., et al. “The Influence of Affective Teacher–Student Relationships on Students’ School Engagement and Achievement.” Review of Educational Research, vol. 81, no. 4, Dec. 2011, pp. 493–529, https://doi.org/https://doi.org/10.3102/0034654311421793

Tarrant, Mark. “Adolescent Peer Groups and Social Identity.” Social Development, vol. 11, no. 1, Jan. 2002, pp. 110–23, https://doi.org/10.1111/1467-9507.00189

Verhoeven, Monique, et al. “The Role of School in Adolescents’ Identity Development. A Literature Review.” Educational Psychology Review, vol. 31, no. 1, Dec. 2019, pp. 35–63, https://doi.org/10.1007/s10648-018-9457-3

Evaluating the Efficacy of a Low-Cost Digital Learning Platform on Academic Performance of Students Aged 5-16 By Ansh Tripathi

Abstract

Low-cost digital platforms are a scalable way to improve academic performance when consistent, teacher-led instruction is limited. This paper asks: how effective are low-cost digital learning platforms in improving academic performance amongst students aged 5 to 16? Using usage data from a Math and English learning platform, effectiveness was assessed using diagnostic-to-final improvement, average scores at attempted levels, and dropout rates, with comparisons by gender (male/female) and age groups (ages 5-9 vs. ages 10-16). Results show measurable gains for each subgroup, with age being the strongest predictor of both performance and persistence. Younger students scored lower, progressed through fewer levels, and dropped out at far higher rates than older students did, but also showed most improvement. Overall findings from the analysis suggest that low-cost digital platforms can be effective at improving performance, especially for learners with lower baselines.

Keywords

Digital learning platforms; academic performance; age-related outcomes; gender-related outcomes; student dropout.

1. Introduction

Access to quality education remains a global challenge, especially in communities facing infrastructure or staffing limitations, which can hinder students' academic progress. In recent years, low-budget digital learning platforms have emerged as a potential solution to bridge this gap. These learning apps are mostly built on basic user requirements for minimal technology, with some requiring internet access rather than continuous teacher supervision. Analytic evidence suggests that these tools aren’t a temporary way of dealing with this problem. A review of technology supporting personalised learning across low- and middle-income users found a significant positive effect on the learning outcomes of students from ages 6 to 15, with math and literacy skills notably improving (Major et al.). Given the increase in global investment in technology and education, it is important to assess not just whether these platforms improve academic performance, but also for which students they work best.

Hence, this paper investigates how effective low-cost digital learning platforms are in improving academic performance among students aged 5 to 16. The “effectiveness” is assessed from the following indicators through data drawn from an interactive, gamified Math and English digital learning platform: (1) the change in student performance between a diagnostic (baseline) quiz and a final quiz, (2) the average student scores across levels attempted, and (3) the number of student attempts per level, or the dropout rate. These indicators are compared across two demographics, those being gender (male vs. female) and age groups (ages 5-9 vs.

ages 10-16). This is done to determine whether the platform's efficacy is consistent across multiple learner populations, and to determine if any population benefits more than another.

Existing research suggests that such digital learning outcomes are rarely uniform across different learning populations. In a study, gender differences have been documented in subjects such as mathematics, where girls using digital math games have achieved greater learning progress than boys despite reporting a lower self-perceived competence level (Cubillos et al.). Additionally, in students learning English as a second language, a large-scale study revealed a gradual shift towards boys outperforming girls, even though the girls continued to lead in basic literacy (Sandsør and Brevik). Similarly, apart from gender differences, age-related differences were well documented, with younger (primary-aged) students reporting lower self-perceived effectiveness of online learning than older students, who have higher levels of self-regulation in digital environments (Zheng et al., 2022). Furthermore, younger students have been found to drop out of online courses at a significantly higher rate than older students (Tan and Li), possibly being a case of course rigor. These findings especially matter to low-cost digital platforms, which rely on students to self-navigate the platform.

Therefore, the purpose of this paper is to assess real usage and performance data from a low-cost digital Math and English learning platform in order to evaluate the overall efficacy, identify the subgroups that benefited the most, and where performance gaps or dropout risks stem from. The scope of this analysis is limited to comparative statistics drawn from the platform's diagnostic, final, attempted level, and dropout data. Each diagnostic quiz provided a score out of 10, while each final quiz had 15 questions. This data was averaged out and converted into percentages. The remainder of the paper is structured as follows: the Results section provides a comparative analysis across Math and English subject outcomes based on gender, and then the same between the age groups (ages 5-9 vs. ages 10-16). Each subsection reports the relevant data, an explanation of the observed pattern, and a discussion of how the findings relate to the wider academic literature on gender-based and age-based online learning.

2. Methodology

2.1 Aims

This study aims to evaluate the effectiveness of a low-cost, interactive, digital Math and English learning platform in improving academic performance amongst students aged 5-16. The objectives of the research are as follows:

- Extract diagnostic, final, attempted-level, and dropout data from the platform's records for both Math and English
- Perform a quantitative performance analysis between male and female students.
- Perform a quantitative performance analysis between younger (ages 5-9) and older (ages 10-16) students

2.2 Research Design

This study uses a descriptive, comparative quantitative design. The analysis relies on data already collected from students' normal usage of the tested platform. This approach was chosen because it reflects authentic student behaviour and performance, instead of fabricating responses based on the fact that the students know they are being surveyed.

2.3 Participants

The participants included in the data set were both male and female students aged 5-16 who used the platform's Math and/or English modules. For data to remain consistent and mitigate any form of error, students were chosen from the same study period.

2.4 Data Collection

Data was collected by exporting the platform's built-in usage and performance analytics for the analyzed period, covering 3 main categories: (1) the change in student performance between diagnostic (scored out of 10) and final assessment (scored out of 15), analyzed through average percentage of correct answers; (2) average score achieved at each attempted level for both subjects; (3) the number of student attempts at each level. This data was then organized and averaged by age groups and gender.

2.5 Data Analysis

The exported data was used to create comparative bar and scatter charts, with trendlines added to visualize directional patterns. Given the descriptive scope of the study, the graphs were analyzed descriptively rather than using inferential statistical testing. For future reference, "N" reflects only students with complete attempt records. Sample sizes vary across analyses due to some students lacking complete records.

2.6 Ethics

This study analyzed anonymized performance data already collected by the platform through students' normal use of the platform. No names or any identifying information were collected or used in the analysis. Data were grouped at the level of gender and age so that no single student's performance can be identified from the results provided.

3. Results and Discussion

3.1 Gender: Comparative Analysis Between Genders

3.1.1 Math

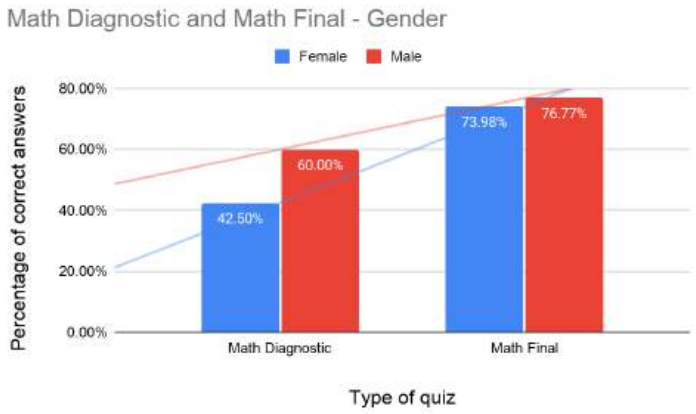


Figure 1: Comparison of Math Diagnostic and Final - Gender (N=35)

Figure 1 shows that on the Math Diagnostic quiz, male students scored 60.00% on average, notably higher than that of female students, who achieved an average score of 42.50%, creating a gap of 17.5 percentage points. By the Math Final, female students reached 73.98% correct compared to 76.77% for male students, narrowing the gap to just 2.79 percentage points. Overall, female students increased by 31.48 percentage points from the diagnostic to the final, almost double the 16.77-point gain achieved by male students. Although the female students began with a noticeable disadvantage, the platform was effective in closing the gap and bringing the final scores of both genders to a similar score on average. Similar findings were found in a study of a digital drill-and-practice math game with primary school students. Despite the female students initially reporting a lower self-perceived knowledge, the study's findings revealed that girls achieved significantly greater learning growth than the boys did over the course of the study (Cubillos et al.). This suggests that structured, repetitive digital environments may help offset an existing gap in self-perception and confidence, as can be seen in both this experiment and the literature, where it produced outsized score improvements for female learners even when their starting point is lower.

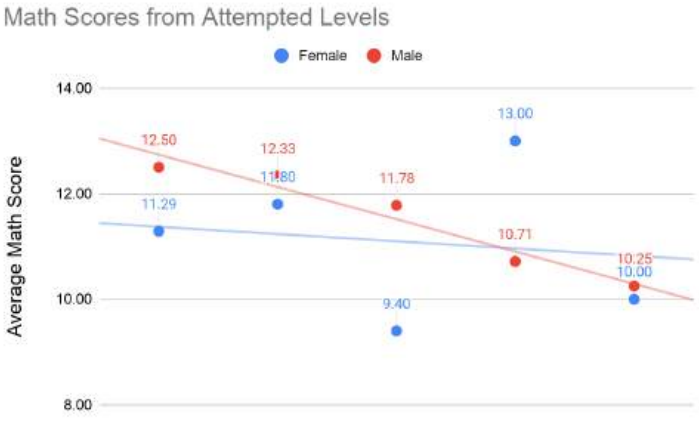


Figure 2: Average Scores of Attempted Levels for Math - Gender (N=35)

Figure 2 shows that male students' average score declines fairly steadily as they progress, falling from 12.50 at level 1 to 10.25 at the last attempted level, that being level 5. Although the overall trendline also slopes gently downwards, ending at a similar point of 10.00, the female students' scores are more volatile across all levels, ranging from a low of 9.40 to a high of 13.00. This indicates that for both genders, the levels attempted later on in the platform are more difficult than the students' ability. Even though the male pattern is more consistent, the female pattern shows drastic swings by each level, likely reflecting the smaller number of females who progressed to a higher level. A study on 206 kindergarten students revealed a similar finding. The study found that girls and boys responded differently as the difficulty of a mathematical pattern task was increased. Their accuracy and error patterns diverged as the difficulty of the task increased, with each gender showing distinct patterns. Specifically, the study found that the girls were significantly more likely than boys to answer correctly as difficulty increased (Schreiber). This offers a plausible explanation for why the average scores by female students in the studied platform here remain flatter in comparison to the trendline for the average male scores.

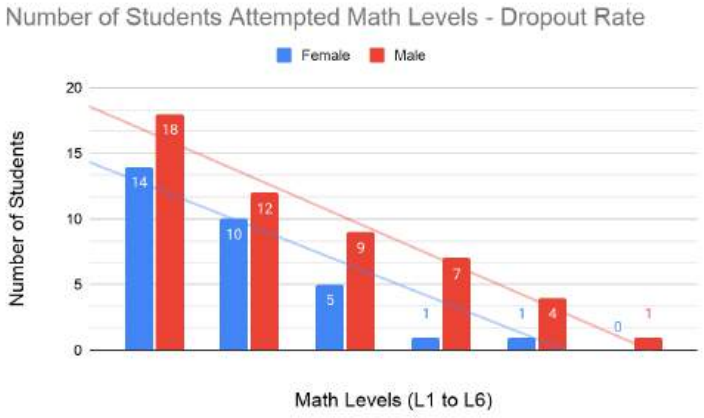


Figure 3: Attempted Levels for Math - Gender (N=32)

Figure 3 reveals a pronounced dropout curve for female students, comparably higher than that of males. While 14 female students attempted Level 1, only 5 reached Level 3, and just 1 student reached Level 4, compared with 18, 9, and 7 male students, respectively, at the same levels. By Level 4, only 1 female student attempted Level 4, which is significantly lower than the 7 male students at the same level. This is a notable finding given that female students, as per Figure 1, showed the larger performance gains. The students making the most academic progress were also likely to disengage from the platform early. This also reveals why the scores for female students, shown in Figure 2, for level 4 were high; there was only 1 female student who attempted level 4, causing the average to increase. This pattern aligns with research showing that female students in digital math environments often report lower self-perceived knowledge than male students, even when their measured learning gains are greater than or equal to those of male students (Cubillos et al.). This may suggest that, aside from ability or performance, female

students' persistence on the platform may also be driven by confidence and engagement factors, meaning low-cost platforms may need built-in encouragement to retain users. However, these factors were not directly measured in the study.

3.1.2 English

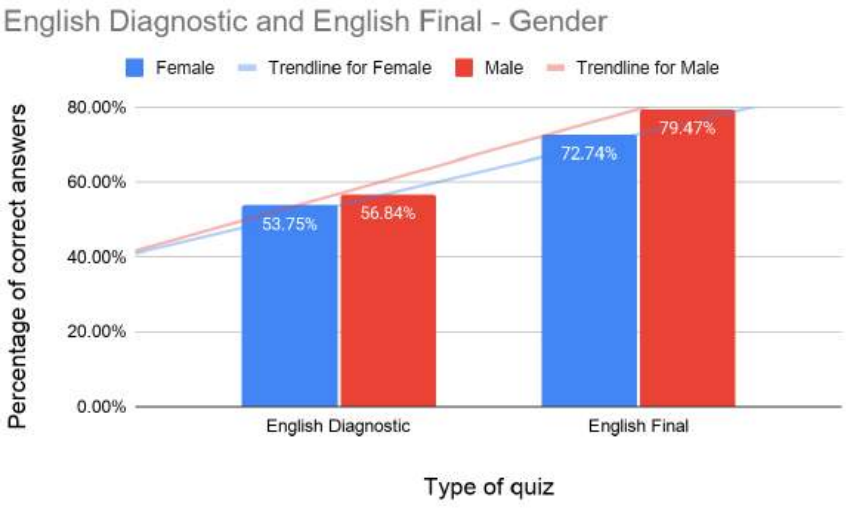


Figure 4: Comparison of English Diagnostic and Final - Gender (N=35)

Figure 4 shows a pattern distinct from the platform's Math results: male students outperformed female students at both the diagnostic stage, where males and females scored 56.84% and 53.75% respectively, and, more markedly, at the final stage, with scores of 79.47% and 72.74% respectively. Unlike in Math, the gender gap is much smaller and consistent. During the English Diagnostic, male students had a higher score of only 3.09 percentage points, which increased to 6.73 points by the final assessment. This finding contradicts the reported gender gap in literacy. As seen in the study done across 63 education systems with five decades of international data, girls typically outperform boys in reading and writing across primary education (Steinmann et al.). However, a large-scale study of students learning English as a second language found that boys performed better in tests on second-language English proficiency, even while girls continued to outperform boys in general during the same period (Sandsør and Brevik). Because this platform's English content is directed towards students who are learning English as a second or sometimes a third language, this reversal of male students pulling ahead over time is consistent with the emerging pattern documented in that literature.

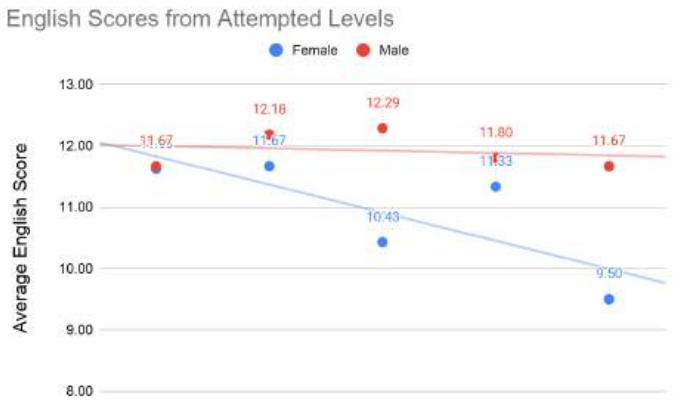


Figure 5: Average Scores of Attempted Levels for English - Gender (N=35)

Figure 5 shows that male students' average English scores remained relatively flat and high across levels, ranging narrowly between 11.67 and 12.29, with a marginal downward trendline. Female students, on the other hand, showed a much steeper downward trendline, falling from 11.67 at the first levels to 9.50 at the final levels. This suggests that as female students progress through the English levels, the material becomes more difficult for them than it does for male students, who are seen to keep a relatively stable performance throughout. This male advantage in students learning English as a second language mirrors a broader shift in L2 English literature. A population-level study of over 1 million Norwegian students from 2007-2018, who learn English as a second language, found that boys have been closing or even reversing the gender gap. Surprisingly, the same female cohort continued to outperform male students in first-language reading assessments in the same period (Sandsør and Brevik). This indicates the score decline seen here may be specific to students learning English as a second language, rather than reflecting a broader literacy weakness amongst female users on the platform. Additionally, the male students' resistance to that decline may be tied to, as the study presumed, compatible, informal technology-based exposure to English outside of the used platform.

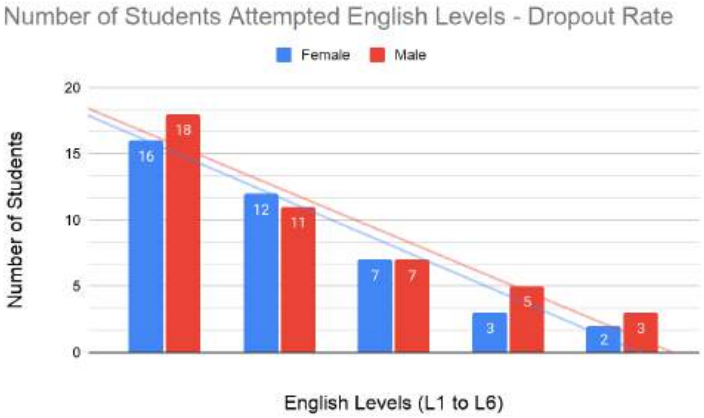


Figure 6: Attempted Levels for English - Gender (N=34)

Figure 6 shows a much smaller gender divergence in English dropout than it does in Math. 16 female students attempted Level 1, with 7 reaching Level 3, and just 2 reaching Level 5, compared to 18, 7, and 3 male students, respectively, reaching and attempting the same levels, revealing a broadly parallel pattern, with the two lines converging by the later levels. This is notably different from the dropout rate observed in Figure 3, where the gender gap is far more pronounced. Research on dropout rates of students working in online environments found that course-related and motivational factors, like course quality, student satisfaction, and academic preparation, are the strongest predictors of whether a student continues or not with a program (Rahmani et al.). This supports the interpretation that the gender-based dropout disparity seen in Math is specific to Math itself, as it is not seen in English, rather than being a generalized tendency of female learners to disengage from this platform.

3.2 Age: Comparative Analysis Between Age Groups

3.2.1 Math

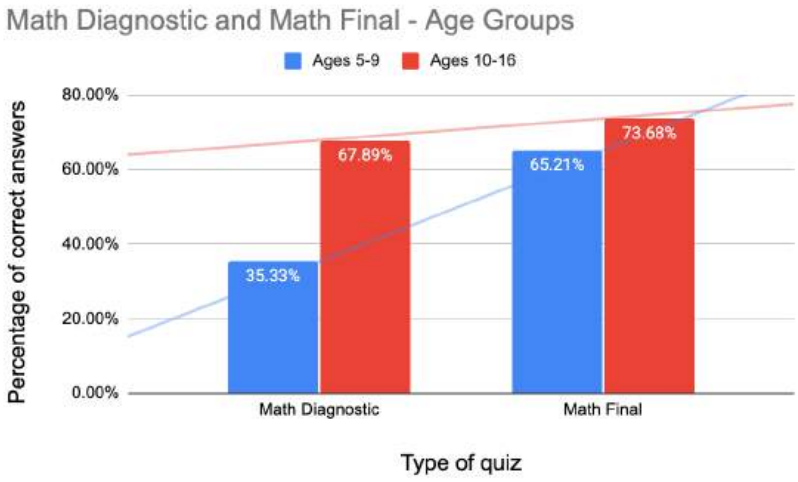


Figure 7: Comparison of Math Diagnostic and Final - Age (N=35)

Figure 7 shows that younger students (ages 5-9) began far behind older students (ages 10-16) on the Math Diagnostic, scoring 35.33% compared to 67.89%. By the Math Final, the younger group had improved substantially to a score of 65.21%, while the older group improved to 73.68%, narrowing the previous gap of 32.56 percentage points to only a gap of 8.47. The younger group's gain of 29.88 percentage points was more than five times that of the older group, which was only a gain of 5.79, suggesting that the platform delivers its largest measurable benefits to the younger learners, even though it doesn't fully close the age-based performance gap. This is broadly consistent with meta-analytic evidence on technology-supported personalised learning in low- and middle-income settings, which found a significant positive effect on learning outcomes across the 6-15 age range, with mathematics and literacy showing similar positive outcomes (Major et al.). The steep improvement amongst younger students here

suggests they may have had more foundational room to grow, making the diagnostic-to-final gain a particularly sensitive indicator of the platform's effectiveness for this age group. Many of the students in the younger group scored relatively worse than students in an older group, indicating a lack of knowledge on the specific topics. After interaction with the platform, the younger students displayed high academic progress, indicating the content was initially new to them.

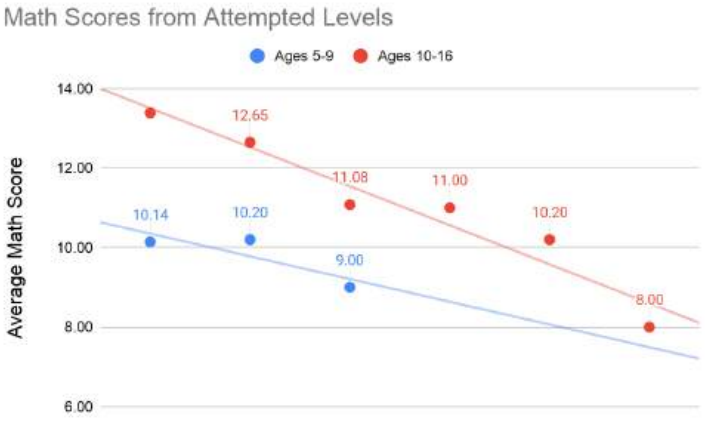


Figure 8: Average Scores of Attempted Levels for Math - Age (N=35)

Figure 8 shows a clear separation of the two age groups. Students in the age group of 10-16 achieved consistently higher average scores across all levels attempted, starting around 13.33 and declining gradually to 8.00 by the final level. On the other hand, students aged 5-9 scored lower throughout, moving between 10.20 and 9.00 across the first three levels. More notably, the older groups' scores span a wider range, reaching all six levels provided by the platform, more than the younger groups, who only reached the first three levels. This matches findings from research measuring digital literacy across age cohorts, which found that older students consistently outperformed younger students, even in content with higher difficulty levels. However, they also found that these older students displayed greater variability in scores as the difficulty of the content and scope increased (Pan et al.). This narrower band of scores in younger students, as seen in Figure 8, likely reflects both their more limited progression through the platform and the more foundational, less variable nature of the content available to them at their attempted levels. However, it also depicts why there was a greater diagnostic-to-final gain, as there were fewer attempts by students in the 5-9 age group, while their scores remained higher at level 3 than the scores achieved by students in the 10-16 age group at level 6.

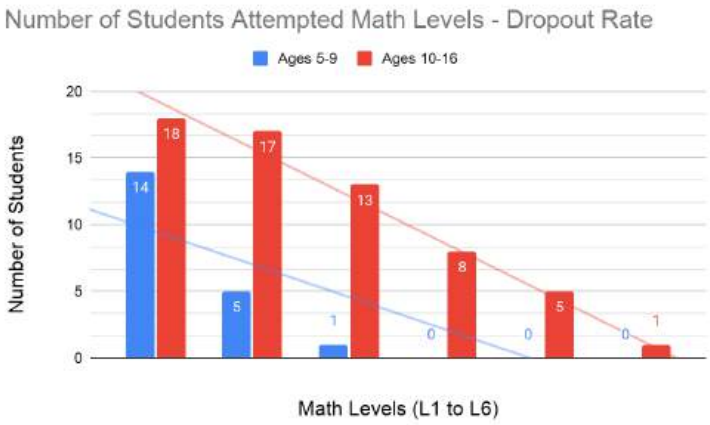


Figure 9: Attempted Levels for Math - Age (N=32)

Figure 9 shows a dramatic difference in persistence between age groups. Among students aged 5-9, participation fell from 14 to 5 to just 1 in Level 3, with no students reaching Levels 4-6. In contrast, students aged 10-16 persisted much further, declining gradually from 18 to 13 to 1, in Levels 1, 3, and 6 respectively. This means that the youngest students in the sample almost universally disengaged from Math within the first two to three levels, while older students continued advancing steadily. This finding directly mirrors that of a large-scale longitudinal study of elementary students in online extracurricular courses, which found that younger students were significantly more likely to drop out of online learning than older students. This was found to be due to less-developed metacognitive skills, weaker self-regulation, and, therefore, a greater difficulty managing independent study without teacher prompting (Tan and Li). For a low-cost platform that depends on independent, self-paced progression, this suggests that younger students may require additional scaffolding to sustain engagement. This also reveals that another reason for the fewer attempts by younger students on the later levels may be attributed to the high course rigor; the findings may indicate that the platform has a lower instructional ceiling for students aged 5-9 than those aged 10-16. One possible explanation is provided by Cognitive Load Theory, which proposes that due to younger students having less-developed prior knowledge and schemas, increasingly difficult levels might impose higher cognitive load (Sweller et al.). Here, younger students may not be able to grasp the content easily, making progression past the first three levels much more difficult.

3.2.2 English

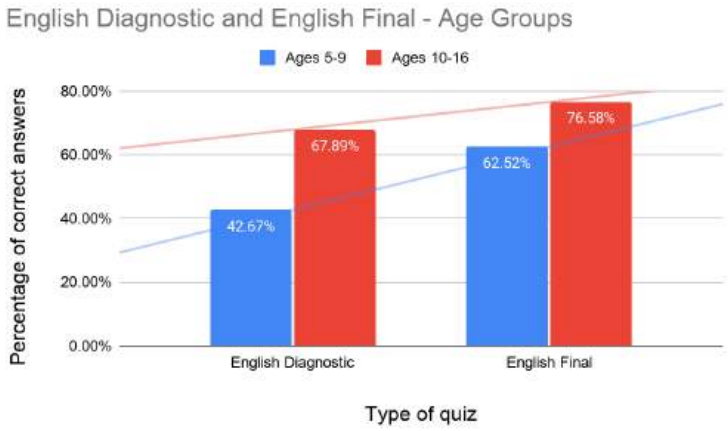


Figure 10: Comparison of English Diagnostic and Final - Age (N=35)

Figure 10 mirrors the same pattern seen in Math. Younger students began far behind older students on the English Diagnostic, scoring 42.67% versus 67.89%, and although they improved considerably to 62.52% by the final assessment, the older group also improved, reaching 76.58% on the English final and leaving a gap of 14.06 percentage points. The younger group's gain of 19.85 points again outpaced the older group's gain of 8.69 points, reinforcing the pattern that the platform delivers its largest improvements to its youngest users. Research into primary students' online learning during remote school found that younger primary students consistently reported lower self-perceived effectiveness of online learning and a greater difficulty in engaging independently as compared to older primary and secondary students, who have more developed self-regulation abilities and confidence with using digital tools (Zheng et al.). Despite their meaningful gains, this helps explain why the younger learners don't fully close the performance gap with the older age group by the final assessment.

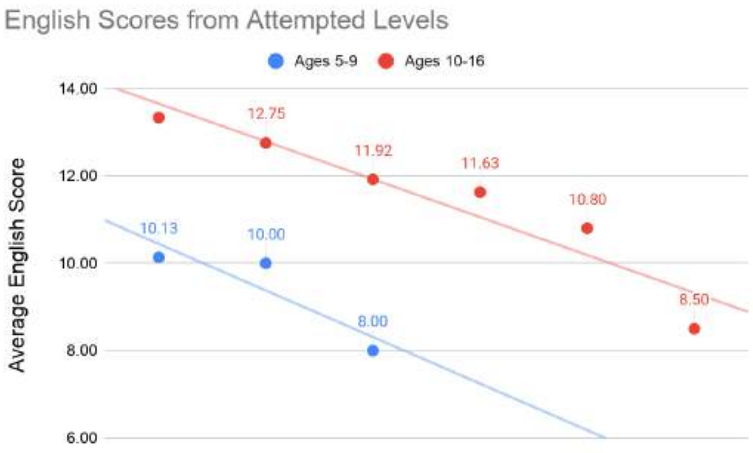


Figure 11: Average Scores of Attempted Levels for English - Age (N=35)

Figure 11 shows that students aged 10-16 sustained relatively high average scores across a long sequence of levels, declining gradually from roughly 13.30 to 8.50 over all of the six levels, while students aged 5-9 declined much more steeply over a short span, from 10.13 down to 8.00 across only three levels before their progress ends. The younger group's steeper and shorter trajectory suggests that English content becomes disproportionately difficult for them more quickly than it does for older students, who are able to sustain performance over a longer stretch of increasingly difficult material. This is consistent with longitudinal evidence that digital and general learning ability improves with grade level, with older students not only scoring higher on average but also engaging with a wider range of content difficulty than younger students, who tend to plateau or decline earlier when independent and self-paced demands increase (Pan et al.). In the case of this learning platform, the younger age group wasn't able to attempt the later levels due to the high complexity level. This reinforces the interpretation that self-regulation and independent comprehension skills, which develop with age, play a meaningful role in sustaining performance on such digital platforms.

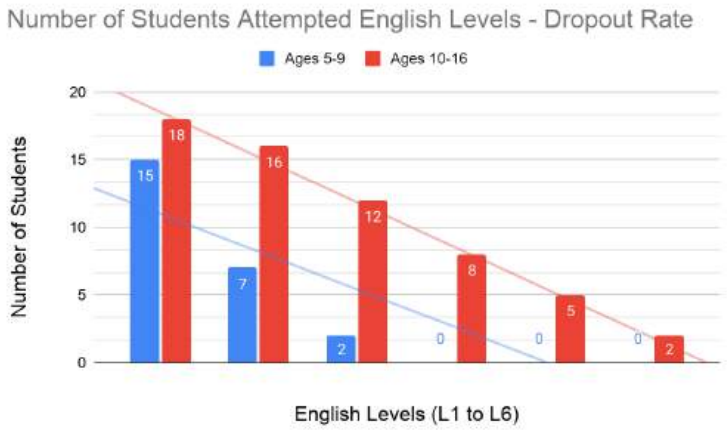


Figure 12: Attempted Levels for English - Age (N=33)

Figure 12 shows the same divergence seen in the Math dropout data. Students aged 5-9 dropped from 15 to 7 to 2 attempts within the first three levels, and none continued to Levels 4 or 5. Students aged 10-16, by comparison, declined gradually from 18 to 16 to 12 to 8 to 5 to 2 across all six levels, representing meaningfully greater persistence at every stage. This age-based attrition gap is well documented in the literature about language learning, and two specific explanations emerge. A longitudinal study tracking 601 Duolingo users over six months found that age was one of the two strongest indicators of whether a learner stopped using the app altogether. In the study, older users were significantly more likely to sustain engagement throughout completion (Sudina et al.). Since that platform, like the one examined here, relies on learners self-navigating content with minimal external accountability, this suggests younger English learners are structurally more vulnerable to disengagement in unsupervised digital environments, despite the inclusion of gamified lessons, because sustaining independent motivation over many levels is itself an age-related skill. The second explanation is about

processing speed rather than motivation, done by the Barcelona Age Factor project, a large-scale study comparing English learners who began at ages 8, 11, 14, and adulthood under equal hours of classroom instruction. They found that students who began at an older age consistently progressed faster and scored higher at every stage measured, with younger learners only catching up, if at all, after a substantially increased exposure (Muñoz). Because this platform delivers English in a similar self-paced format, younger students may simply be encountering material that is harder for them to process, effectively making their early dropout rate a result of a pace of visible progress too slow to sustain engagement, rather than being a lack of persistence itself. Together, these findings reinforce the pattern across Figures 9 and 12; for low-cost platforms that depend on self-paced progression, age-related persistence and processing speed, rather than subject-specific difficulty alone, may be the dominant factors separating students who complete the program from those who abandon it early.

4. Conclusion

4.1 Diagnostic to Final Assessment Improvements

Across both demographics tested, improvement was overall consistent. The group that started further and scored less on the diagnostic had a greater improvement by the final assessment for each level. Female students gained more than male students in Math (31.48 vs. 16.77 points), while the reverse held in English, where by the final assessment, male students started slightly ahead and pulled further ahead. On the other hand, age showed a more uniform pattern, where, in both subjects, the 5-9 groups' gains were roughly three to five times larger than that of the 10-16 group, with a residual gap remaining at the final assessment, despite the catch-up growth. Taken together, this suggests that diagnostic-to-final improvement is also shaped by the starting point rather than by gender and age groups alone. The direction of the gender effect is subject-dependent, favouring female students in Math and male students in English, whereas the direction of the age effect is consistent across both subjects. Younger groups are favoured in terms of raw gain, but don't fully close the age gap itself due to their ceiling being lower than that of students in higher age groups.

4.2 Average Scores at Attempted Levels and Dropout Data

Age produced a far larger and more consistent split than gender did when observing scores students achieved at attempted levels. Older students (10-16) scored higher on average than younger students (5-9) in both Math and English, and progressed through a longer sequence of levels overall. The younger students' data stopped after three to four levels, which also contributed to their lower average scores. It also indicates a genuine ceiling: most 5-9-year-olds simply never reached the harder content that would test them further. The youngest students appear to hit a hard ceiling in both subjects at roughly the same point in the sequence, consistent with age-related limits on independent self-regulation rather than anything specific to Math or English, speaking to the course rigor of the platform. Gender, on the other hand, didn't produce

the same structural difference in how far students progressed. Here, differences were more subject-specific than due to gender. This suggests that age, more than gender, determines how much of the platform's content a student is even able to demonstrate their ability on.

This paper is set out to answer the question: how effective are low-cost digital learning platforms in improving academic performance amongst students aged 5-16? As shown above, students demonstrated measurable improvement following their use of the platform for every subgroup examined, but that effectiveness was shaped far more consistently by age than by gender. The practical implication is that low-cost platforms serving a wide age range may need age-differentiated design, such as shorter levels, more frequent feedback, or light external check-ins.

This study was limited to descriptive comparative analysis of one platform’s data, and didn’t test causal mechanisms behind the patterns observed. Future research could track the same students longitudinally to see whether dropout rates amongst younger students remain permanent, and survey the disengaged younger students directly to determine whether dropout stems from difficulty, boredom, or lack of support to further improve the platform. Overall, the data from the platform show evidence of improved academic performance after platform use; although the descriptive design does not establish causation, it does show genuine overall improvement across all demographics in English and Math.

Works Cited

Cubillos, Claudio, et al. “A Digital Math Game and Multiple-Try Use with Primary Students: A Sex Analysis on Motivation and Learning.” Behavioral Sciences, 2024, https://doi.org/10.3390/bs14060488.

Major, Lee, and Gill Francis. “The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis.” British Journal of Educational Technology, 2021, https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13116.

Muñoz, Carmen. “The effects of age on foreign language learning: The BAF project.” Age and the Rate of Foreign Language Learning, 2006, https://www.researchgate.net/publication/292251866_The_effects_of_age_on_foreign_language_learning_The_BAF_project.

Pan, Qianqian, et al. “Measuring digital literacy across ages and over time: Development and validation of a performance-based assessment.” Education and Information Technologies, 2025, https://link.springer.com/article/10.1007/s10639-025-13592-8.

Rahmani, Amir Mohammad, et al. “Dropout in online higher education: a systematic literature review.” International Journal of Educational Technology in Higher Education, 2024, https://doi.org/10.1186/s41239-024-00450-9.

Sandsør, Astrid M. J., and Lisbeth M. Brevik. “Gender gaps in English as a second language over time: a shift in favour of boys.” Assessment in Education: Principles, Policy & Practice, 2024, https://doi.org/10.1080/0969594X.2024.2423151.

Schreiber, Iris. “Gender differences in early childhood mathematics: boys’ and girls’ responses to changing pattern task difficulty.” Cogent Education, 2025, https://doi.org/10.1080/2331186X.2025.2486484.

Steinmann, Isa, et al. “International Reading Gaps between Boys and Girls, 1970–2016.” Comparative Education Review, 2023, https://doi.org/10.1086/724089.

Sudina, Ekaterina, et al. “L2 grit and age as predictors of attrition in mobile-assisted language learning.” Learning and Individual Differences, 2025, https://doi.org/10.1016/j.lindif.2025.102704.

Sweller, John, et al. “Cognitive Load Theory.” Springer Nature Link, Springer, 2011, https://link.springer.com/book/10.1007/978-1-4419-8126-4.

Tan, Haotian, and Xueting Li. “Dropout in Online Education: A Longitudinal Multilevel Analysis of Elementary Students’ Extracurricular English Course Engagement in China.” Behavioral Sciences, 2025, https://doi.org/10.3390/bs15040483.

Zheng, Xiaoxiang, et al. “Primary School Students’ Online Learning During Coronavirus Disease 2019: Factors Associated With Satisfaction, Perceived Effectiveness, and Preference.” Frontiers in Psychology, 2022, https://doi.org/10.3389/fpsyg.2022.784826.

Micro-Activism, Algorithms, and Youth Civic Engagement in the Age of Short-Form Video: A Review of the Literature By Trisha Mittal

Abstract

This literature review studies the evolving landscape of youth civic engagement in the digital era, focusing on the rise of short-form video and algorithmic curation. It explores key themes including digital networked participation and “microactivism”, arguing that actions such as liking, reposting, and commenting serve as legitimate methods for Gen Z and Gen Alpha to develop civic identities. The review analyzes how recommendation algorithms create a reciprocal loop of engagement that exposes youth to diverse causes while simultaneously posing risks such as "issue bubbles," misinformation, and emotional distress. Additionally, it reviews the influence of content creators, noting a shift in trust toward "authentic" digital voices over traditional institutions. A central focus is placed on the online-to-offline transfer, establishing that digital participation is a meaningful indicator of future offline behavior, such as voting or volunteering. The paper addresses critical concerns regarding equity, digital divides, and "compassion fatigue" resulting from constant exposure to global crises. It concludes that to support a new generation in the digital age, society must bridge the "commitment gap" through modernized civic media literacy and tools that translate online awareness into sustained, tangible action.

Keywords youth civic engagement, digital activism, micro-activism, Gen Z

I. Introduction: Why Youth Civic Engagement Looks Different Now

How Has the Meaning of “Civic Engagement” Changed for Gen Z and Gen Alpha?

The meaning of civil engagement has expanded beyond traditional civic acts such as voting and moved to include digital and social activism. This definition of civil engagement includes attending protests, signing petitions, calling representatives, and advocating on social media platforms. As the first generation to have access to social media at an early age, Gen Z stays engaged by posting about social and political issues online such as Black Lives Matter, #MeToo, and reproductive rights. Gen Z’s identity as a highly diverse generation—48% identifying as people of color and 1 in 4 as LGBTQ—is also intertwined with their civic engagement as it serves as a response to first hand experiences (Knifton et al., 2024).

Should Liking, Reposting, Commenting, or Sharing Civic Content Count as Civic Engagement?

Recent scholarship has defined civic engagement to include digital networked participation (DNP) where digital acts such as liking, commenting, and sharing are legitimate forms of social engagement (Wang & Sundar, 2022). Though these actions are less resource demanding than traditional activism, they still have the effects of encouraging action: mobilizing

others to vote or volunteer, signing petitions, influencing others. As social media involvement plays a huge role in youth lives, engaging in online political and civic networks are crucial to how they are socialized into politics. Though these actions count as engagement, there are negative consequences as social media algorithms can push biased posts increasing polarization in comparison to other forms of engagement (Sormanen et al., 2025).

What Is the Difference Between Political Engagement, Civic Engagement, Activism, and Social Awareness?

Civic engagement is multifaceted and a broader term used to describe behaviors where individuals contribute to society. This includes traditional political engagement within institutional political structures: voting, joining political organizations, contacting representatives; however, it also includes community based engagement outside of these structures: volunteering at community events, protesting, civil disobedience. Political engagement works within political structures, while activism works outside of it. Social awareness is a raised cognitive awareness and critique of oppressive systems that perpetuate injustice. This is more heavily linked to activism and it serves as the framework for the way youth understand how specific issues interact with existing systems of power. For Gen Z college students these lines between political engagement and activism are increasingly blurred; for example, organizing against voter suppression focuses both on voting and challenging systemic racial injustice (Kornbluh et al., 2022).

Why Are Young People Especially Influenced by Digital Spaces When Forming Civic Identities?

Younger individuals are especially influenced by digital spaces when forming civic identities as digital spaces provide a constantly accessible environment for self-expression. Social media provides 24/7 access to diverse voices and political opinions; young people also get a chance to express their own narratives that reflect their experiences. Additionally, teenagers may also turn to digital spaces because of the lack of political education in schools which often avoid controversial and current topics. Digital spaces also provide a form of pressure forcing younger people to make opinions on nuanced issues with a focus on being morally correct. This digital environment can create deep levels of political polarization which especially influences younger people ages 16–24 (Stahl & Hu, 2025).

What Makes Short-Form Video Different From Older Forms of Online Activism, Such as Blogs, Facebook Groups, or Long-Form YouTube Videos?

Short-form video content is different from older forms of online activism because it makes content more palatable and shareable especially for younger audiences. Short-term video content like TikTok and Instagram Reels are built on memes and viral audios. This allows for the information to spread quickly and easily, but it can also allow for misinformation and trivialize important situations. Additionally the algorithm of short-term video content allows content to be

pushed across millions and ideas are presented efficiently. Short-term video-content gives a human perspective where younger generations can see perspectives that include them; this is in contrast to Facebook groups or blogs traditionally intended for older audiences (Hautea et.al., 2021).

II. Background: Youth, Social Media, and the New Civic Environment

Which Platforms Are Most Commonly Used by Teenagers Today?

As of 2025, YouTube is most commonly used with 9 in 10 teenagers among the 15–17 and 13–14 age group reporting they have used it. TikTok and Instagram follow after with 6 in 10 teenagers using it. 75% of teens use YouTube daily, 61% use TikTok daily, and 55% use Instagram daily. One in five teens report using YouTube and TikTok almost all the time (Pew Research Center, 2025).

Why Are TikTok, Reels, and Shorts Especially Powerful for Civic Messaging?

TikTok, Reels, and Shorts are especially powerful for civic messaging because they allow for viral audios and concepts which can be remixed onto a diverse population. This can spread civic messaging faster than ever before without having to be searched for, and without having to read full documents or 30 minute long YouTube videos. They make heavy topics more palatable for all audiences and make civic discussion and messaging feel more authentic and relatable (Yu et al., 2025).

How Does the "For You Page" or Recommendation Algorithm Expose Teens to Social Issues They May Not Have Searched For?

Recommendation algorithms expose teens to social issues they may not have searched for by reaching content outside of the user's typical social circle. They push viral topics and teens are exposed to a greater percent of diversity and experiences they wouldn't have encountered before. Recommendation algorithms are designed to maximize engagement (likes, follows, shares) by pushing out content to many; this is especially evident with controversial social issues as they can garner the most amount of support/anger, in turn, boosting engagement (Metzler & Garcia, 2024).

Are Young People Actively Seeking Civic Information, or Are They Passively Encountering It While Scrolling?

Research shows that there is a reciprocal loop when it comes to how young people encounter civic information online. Rather than a choice between one or the other, most young people passively find civic information through scrolling, and this exposure often leads to political behaviors which further reinforce political content and civic information online. Even when users don't intentionally seek out this content, algorithms track their engagement and behavior and serve them similar content. In a 2016 study, participants rated their frequency of

coming across news online incidentally at a mean of 2.61 to 2.74 on a 4 point scale. And in a 2018 study it was found that there is a significant positive relationship between being active and how much civic information a person encounters. There is a reciprocal loop where young people find civic content passively, engage with it, and then continue to be fed civic information by online algorithms (Lee & Xenos, 2022).

How Does Constant Exposure to Global Problems Affect Youth Civic Identity?

While constant exposure to global problems can cause stress and uncertainty within the youth, it is the primary driver for youth civic identity to explore new beliefs and create a more mature identity. Adolescents find their identity by questioning the status quo and their place in society; awareness of stressful divisive conditions force them to reconsider their civic identity and explore. Constant exposure to global problems greatly extends periods of exploration where younger individuals find political identities that align with their values. This can also strengthen pre-existing civic beliefs, especially beliefs that expand the wellbeing of individuals' culture and communities (Branje et al., 2021).

Does Social Media Make Civic Engagement More Accessible for Students Who Cannot Attend Protests, Join Organizations, or Vote Yet?

Social media significantly increases the accessibility of civic engagement for students by lowering barriers such as time, commitment, and eligibility restrictions. For students who cannot yet vote, social media serves as a tool for political socialization. Students can use digital platforms to share civic information, sign petitions, boycott corporations, and learn about various issues. It serves as a tool for young people to be engaged with their community and encourage others to take action. In a meta analysis of 36 studies, 80% of coefficients showed a positive relationship between political participation and social media use (Lee & Xenos, 2022).

III. Defining "Micro-Activism"

What Counts as Micro-Activism?

Micro-activism can be social media engagement like resharing social justice posts, sharing information, platforming causes, signing petitions. It can also be small changes in behavior that challenge different issues such as recycling, boycotting, and using public transport. Micro-activism can also be creative acts like writing letters to policymakers, donating to fundraisers, volunteering, or using art to express civic ideas (Hughes, n.d.).

Is Micro-Activism Meaningful, or Is It Mostly Symbolic?

Although micro-activism can appear symbolic, it is very meaningful for younger individuals and those who do not have access to bigger activities; it gives a chance for every individual to have a voice and provides tools to challenge current systems. Sustained digital activism can become meaningful as it brings attention to issues and advocates for change. In a

Cabotagestudien study, a decade-long project of crowdsourcing and social media helped to collect data which led to industry wide policy changes in the EU. This shows how micro-activism may not be as long-lived as stronger acts like voting, but it still remains a powerful tool to spark change, all while remaining inclusive. Younger individuals, those with disabilities, financial restraints, illnesses, and other setbacks all have a chance to be changemakers (Mindel et al., 2024).

Can Small Digital Actions Create Real Awareness, Pressure, or Fundraising?

Small digital actions can be highly effective at raising awareness, pressure, and fundraising and serve as an entry point for meaningful civic engagement. The internet is best at spreading awareness and bringing attention to issues that may often be overlooked. This is evident, for example, with young Palestinians sharing their perspectives on social platforms like TikTok; though it may seem insignificant, interacting with these posts has restored many individuals' livelihoods and brought them into safer conditions. Small digital actions mobilize greater amounts of people as they are easier actions to do, helping many practice civic engagement. Online platforms are also very effective in fundraising, helping donation requests reach a massive audience. Examples of this are the ALS Ice Bucket challenge which raised over 100 million dollars for research and the International Red Cross which raised 7 million in four days for the 2010 Haiti earthquake. These fundraisers were low-cost and low-effort to participate in. Yet, they raised millions of dollars and impacted millions of lives, demonstrating the effectiveness of small digital actions in fundraising and creating real awareness/pressure (Cooper, 2023).

When Does Micro-Activism Become "Slacktivism"?

"Slacktivism" refers to performative online activism that rarely results in tangible change. Micro-activism is distinguished from "slacktivism" based on its impact and relationship to effortful forms of civic engagement. Slacktivism hinders real change only when it is treated as a trade off from more effortful action; this can be seen, for example, when an individual simply likes a civic post and believes their quota for civic engagement is met, leading them to "trade-off" effortful action. Micro-activism provides a foot in the door for other forms of activism but it should be seen as "enough" to ignite change. It is not a replacement for more significant forms of activism, but rather an addition. Micro-activism in any quantity will always be more impactful than no activism, which makes "slacktivism" a simplistic view of online activism. The efforts from micro-activism are significant as online and offline activism are intertwined and positively related (Greijdanus et.al., 2020).

Does Calling It "Slacktivism" Unfairly Dismiss Youth Participation?

Calling micro-activism "slacktivism" dismisses youth participation and ignores the privilege of accessibility to traditional participation platforms. Much of Gen Z and Gen Alpha are under voting age, yet are still impacted by local bills and policies. Online activism provides

an accessible gateway for collective civil action. The “slacktivism” label fails to recognize social media’s vast capabilities in mobilization. Youth participation is helpful in any amount whether that be more effortful activism or simply sharing a post. “Digital Dualism”—the idea that online and offline realities are separate—is an ignorant perspective which fails to account for the deep intertwined nature of the digital world and our everyday lives. Micro-activism has led adolescents to create significant tangible civic change and should not be dismissed as “slacktivism” (Greijdanus et.al., 2020).

Are Low-Effort Actions Still Valuable if They Are the First Step Toward Deeper Engagement?

Low effort actions are valuable because they are the first step towards deeper engagement. Sustainable behavioral and structural change cannot come without low-effort actions to support it and pave the way. Small achievable behaviors reduce the resistance to change by breaking down large civic tasks into manageable ones. Significant commitments to political activism can be impractical and overwhelm individuals in a way low-effort activism does not. The principle of marginal gains states that a 1% improvement in behavior can lead to significant cumulative benefits; low-effort actions compound and strengthen the success of deeper engagement (Akash & Chowdhury, 2025).

How Do Young People Themselves Describe These Actions: Activism, Awareness, Support, Trend-Following, or Social Pressure?

Young people define activism by its intent to bring systemic change. Youth spaces have created a distinction between genuine activism (online education and real effort) and performative activism (acts for social attention). Youth perspectives have shifted the definition of awareness from low-effort activism to high effort. Raising awareness requires knowledge of a topic and creating content to educate masses. It is an essential part of advocating for change as change cannot come before awareness. Younger individuals also advocate for support through life-style changes rather than traditional civic acts. This can include boycotting, recycling, ethical consumption and a “shift” in culture. Digital environments have also cultivated a social pressure economy where there is pressure for the youth to pick a side on complex nuanced issues causing the youth to follow trends and adopt positions before they are fully informed (Oinonen & Paloniemi, 2023).

IV. Algorithms and Civic Awareness

How Do Algorithms Decide Which Civic Issues Become Visible to Young Users?

Algorithms decide what civic issues become visible to young users through machine learning techniques designed to maximize screen time and engagement. Algorithms analyze data such as age, gender, browsing history, location, and purchasing habits in order to categorize user behavior. Platforms use natural language processing, linear regression, and clustering to predict

what types of civil content will keep the user engaged. Algorithms are also designed to push content that triggers the mesolimbic system (brain's reward system). This means issues and content with high-arousal levels that trigger emotions receive greater visibility. Both these processes create a feedback loop between personalized data collection and a user's biological reward centers to push certain civic issues more heavily than others (De et al., 2025).

Why Do Some Issues Go Viral While Others Remain Invisible?

Issues go viral when they evoke heightened high arousal emotions that push people to interact. Content that elicit emotions such as awe, anger, or amusement lead to significantly higher sharing rates than content that is calm or "boring." Issues also go viral based on timing and culture; content that has immediate relevance creates a personal connection which drives engagement. Issues remain invisible when they lack emotional intensity and fail to connect with the general population (Gentry, 2025).

Do Algorithms Reward Simplified or Emotionally Intense Civic Content Over Nuanced Explanations?

Algorithms reward emotionally intensive, simplified, and partisan civic content over nuanced and reflective explanations. Platforms aim to maximize the time spent and engagement of videos; this means that algorithms favor content that captures attention quickly. Algorithms personalize feeds by predicting what a user will engage with in the moment rather than what they would engage with after reflection. This is evident in a study of Twitter's algorithm which found that users did not prefer the partisan political tweets the algorithm selected for them, showing that engagement metrics capture impulsive behavior over content that users find reflective and valuable. Research has found that algorithms exploit human social learning biases by pushing what is known as "PRIME" content; this content is in-group focused and emotional. Platforms push content that amplifies negative emotions, primarily anger. Content that angers the public is significantly more partisan and emotional than nuanced political content. Nuanced explanations require deliberate thought and time while simplified emotionally charged explanations are fast and intuitive. Nuanced explanations fail to align with users' short attention span and therefore are disadvantaged by the algorithm which prioritizes immediate views and engagement (Milli et al., 2025).

How Can Algorithmic Feeds Create "Issue Bubbles," Where Teens Repeatedly See One Side of an Issue?

Algorithmic feeds create "issue bubbles" where teens repeatedly see one side of an issue through advanced machine learning which tailor content to match an individual's preferences. This takes advantage of clustering and continuous feeds which prioritize personalization. Platforms analyze personal information to predict user interaction. To ensure high engagement, algorithms push targeted content which often include biased content and extremist material. In 2022, children and adolescents in the U.S. generated $11 billion in advertising revenue on social

media. This highlights the significant economic advantages of keeping teens in high engagement "issue bubbles" (De et al., 2025).

Can Algorithmic Exposure Help Students Discover Causes They Would Not Learn About in School or at Home?

Algorithmic exposure can help students discover causes and social issues they might not encounter in traditional environments like school or at home. Algorithms prioritize content over common social networks which allows civic messages to reach students regardless of their social circles and prior interests. Although algorithms are rooted in past interactions, they also consider location and viral topics that capture attention. These algorithms can highlight relevant stories and causes; this is particularly valuable for Gen Z who largely depend on social media to stay informed. Digital platforms and algorithmic exposure can fill the gaps left by schools on divisive issues (Swart, 2021).

What Are the Risks of Youth Forming Civic Opinions From Short, Emotionally Charged Videos?

Youth forming opinions from short emotionally charged videos can affect them mentally and increases uninformed decision making. A recent study by Professor Alsharawy found that young adults who consume daily content from political TikToks reported higher levels of feeling distressed, afraid, and upset. This can affect the emotional well-being of adolescents, especially when civic content is intentionally exaggerated and emotionally charged. These videos may also not contribute to a more informed and evolved identity. In this same study, while there was an increase in negative emotions, there were no significant changes in political opinion. As adolescents are undergoing cognitive development, constant exposure to content designed for emotional arousal can cause shifts in the brain region responsible for decision making. This can make it hard for the youth to critically evaluate extremist biased perspectives demonstrated by many short form political videos (UTHealth Houston School of Public Health, 2025).

V. Influencers, Creators, and Peer Networks

Why Might Young People Trust Creators More Than Politicians, Teachers, or News Organizations?

Young people may trust content creators more than politicians and organizations due to their personal reliability and perceived transparency. By sharing their lives online, content creators feel authentic and trustworthy; they demonstrate values based on their lived experience, unlike faceless organizations and out of touch politicians. This perceived reliableness gives them credibility. Their voices and experiences feel more honest than dry textbook facts. In contrast to traditional journalism which emphasizes neutrality, young adults increasingly prefer creators who are open about their views and biases. In a study by Pew Research Center, 63% of adults between the ages of 18–29 believe it is acceptable for journalists to advocate for the communities

they cover, on social media. Additionally 50% of young adults have a lot/some trust in social media organizations, trusting social media at a higher percent than national news. Young adults prefer transparency and the relatable nature of content creators which may be the reason young people lean towards social media creators over politicians and organizations (Pew Research Center, 2025).

How Do Influencers Make Civic Issues Feel More Personal or Understandable?

Influencers make civic issues feel more personal and understandable through emotional "authentic" storytelling and simplified messaging. Influencers create a "curated intimacy" with viewers which can foster into an imagined close bond called a parasocial relationship. This perceived authenticness and closeness can make their civic messaging feel less formal and increase the audience's trust. It can lower resistance to hear civic issues and information as their advice feels more like it's coming from a friend rather than a formal figure. Influencers use emotional simplified messaging that is easier to comprehend while scrolling online. They make distant civic issues feel more personal by creating emotional, simplified, and "authentic" content (Klüser, 2025).

What Happens When Civic Education Is Mixed With Entertainment?

When civic education is mixed with entertainment, it blurs boundaries between political information and leisure. It provides accessibility in reaching a more diverse, younger audience and allows civic education to surpass traditional institutional resources. The variety of formats within entertainment (long-form videos, short-form videos, social media posts, music) combined with the variety and diversity of creators increases the audience of this content significantly in comparison to news articles (Schmitt et al., 2024).

Are Lifestyle Creators Who Occasionally Discuss Social Issues More Persuasive Than Openly Political Creators?

Lifestyle creators who occasionally discuss social issues can be more persuasive than openly political creators. This is because of these creators' perceived authenticity, as mentioned earlier. When creators who do not typically operate within the political sphere participate in civic discourse, their voices are seen as more genuine and independent. Lifestyle influencers are invested into appearing reliable and relatable to project a lifestyle their followers admire; this creates a more emotionally intimate relationship between the creator and the viewer, making these individuals' values feel more trustworthy. Their political commentary feels less coordinated or like an advertisement than a political content creator who focuses entirely on civic issues; therefore, lifestyle creator's voices can actually be much more persuasive than political creators (Stahl, 2025).

How Do Peer Reposts Affect Whether a Teen Sees an Issue as Important?

Peer reposts impact how a teen perceives social and civic issues by acting as an identity signal which enforces social norms. Teens publicly reposting different issues increase the pressure in their peers to also "pick a side" on civic issues. According to a study by Hartl, the odds of friendship decrease by 20–80% for every standard deviation difference in opinion/attribute. Because adolescents are afraid of the social consequences of non-conformity, they often adopt the same civic concerns as those who repost them—changing the way they perceive the severity of different issues and creating pressure to participate (Laursen & Veenstra, 2021).

Can Creator-Led Activism Oversimplify Complex Social or Political Problems?

Although creator-led activism is currently on the rise, it can certainly oversimplify complex social and political problems. Creator-led activism, specifically on social media platforms, are often small bits of information carried through viral sounds and trends; this means that this information can be click-baity and not thoroughly nuanced or researched. Additionally, creator-led activism often fosters specific audiences which can manifest into echo chambers which fail to analyze various perspectives. Due to the heavy misinformation circulating on the internet in these spaces, social media often polarizes issues and simplifies complex problems (Rachimoellah et al., 2024).

VI. Digital Literacy, Misinformation, and Civic Confusion

How Do Teens Know Whether Civic Content Online Is Accurate?

Teens can often be misguided on whether civic content is accurate. Adolescents are often persuaded easily by the appearance and content of a post, rather than its source. This can negatively impact the way they perceive different content and can misguide them in finding accurate civic information. In one study by the Stanford History Education Group, it was found that more than 80% of middle school students believed that sponsored content was a real news story. On the high school level, on a post about President Trump, only a quarter of the students could identify the purpose of the verification checkmark and over 30% argued a fake account was more trustworthy due to its graphic elements. These statistics show that teenagers are often misguided on real and fake content online (Stanford Graduate School of Education, 2016).

What Makes Misinformation Spread Quickly Among Youth?

Misinformation spreads quickly among youth because of a lack of education on source verification and manifests significantly within the current era of click-baity social media posts and artificial intelligence. Education materials on web credibility in schools are often outdated; researchers argue that many schools use tools from 1999 which have not evolved to include the current complexities of the online world. With the AI boom of the 2020s, fake pictures, videos, posts, and articles can be created instantly catapulting the amount of misinformation spread online. This information is often spread through click-baity posts on social media used to create

virality, often targeting the youth as their audience (Stanford Graduate School of Education, 2016).

What Role Should Schools Play in Teaching Civic Media Literacy?

Schools should play a significant role in teaching media literacy. Low levels of media literacy among youth are directly linked to political apathy. Structured programs in schools can change this by building political efficacy—the belief that an individual can understand and influence politics. Schools must teach students the tools to identify propaganda and misinformation, analyze the credibility of sources, and critically evaluate government policies (Muringa & Adjin-Tettey, 2024).

How Can Students Learn to Check Sources Without Becoming Cynical or Disengaged?

Students can learn to evaluate sources without becoming disengaged by adopting lateral reading. Lateral reading is a skill used by professionals to verify and contextualize online content. Instead of looking at a site's aesthetics to determine its credibility, lateral reading involves opening different websites to investigate and fact-check sources. The Digital Polarization Initiative (DPI) simplifies media literacy into four strategies: searching for credible sources, finding the original source, researching the author/organization, and retrying different search paths. These strategies can give students the confidence and tools to navigate civic information online without discouraging or misleading them (Brodsky et al., 2021).

Does Misinformation Reduce Civic Engagement, or Does It Redirect Engagement Toward Poorly Understood Causes?

Misinformation does not typically reduce civic engagement, instead it mobilizes and redirects engagement towards polarizing ideas and factually incorrect causes. Fake news is a powerful tool for online political mobilization. In the 2016 U.S. election, BuzzFeed found that false stories from fake, partisan sites received more engagement than legitimate sources and stories did. This misinformation results in real-world mobilization which hinders the democratic process preventing citizens from considering different points of view and making educated decisions (Olaniran & Williams, 2020).

VII. From Online Awareness to Offline Action

What Types of Online Civic Engagement Are Most Likely to Lead to Offline Action?

Interactive engagements online are most likely to lead to offline civic action. These activities can include posting, sharing, or commenting on civic posts rather than passively watching and scrolling. Offline action is most likely to occur when users are both producers and consumers in civic networks; this pushes users to advocate for their ideas and spark conversations about civic matters (Park et al., 2023).

Does Reposting a Cause Make Someone More Likely to Volunteer, Donate, Protest, Vote, or Contact Officials?

Digital actions like reposting civic issues serve as stepping stones for more demanding traditional acts of civic engagement. Online participation is a significant indicator for an individual's choice to volunteer, advocate, donate, protest, or vote for a certain cause. According to a study done by Galais and Anduiza, donating online creates a 28% increase in joining a civic demonstration as well as a 22% increase in joining a strike and sharing political opinions via email creates a 6% increase in the likelihood to vote. This data was collected longitudinally 6 to 12 months later proving a change in offline behavior after online actions (Galais & Anduiza, n.d.).

What Role Do Schools, Clubs, Nonprofits, and Community Organizations Play in Converting Online Interest Into Action?

Schools, clubs, nonprofits, and community organizations connect a bridge in converting online interest to action. Nonprofits and community organizations, specifically, mobilize individuals by connecting them with causes they align with. Online platforms like Facebook offer targeted advertising formats which organizations can use to their advantage. Research from Boulianne shows that following a nonprofit on Facebook triples the odds of volunteering and donating money offline. Schools and clubs also function to bring people of a certain cause together and direct them to resources promoting action (Boulianne, 2022).

Are Youth More Likely to Act Offline When Digital Content Gives Them a Clear Next Step?

Youth are more likely to act offline when digital content gives them a clear next step; it provides a path of direction to create real-world change. Digital content and engagement channel emotions and often provide calls to action. These "calls" direct individuals on what steps to take and how to channel their emotions into direct actions. For example, undocumented youth in the "DREAMer" movement used online video testimonials to organize direct actions at the government office which resulted in real policy changes like the DACA (Deferred Action for Childhood Arrivals) program—an immigration policy established by the Obama administration in 2012 (UNICEF Office of Research – Innocenti, 2020).

What Barriers Stop Students From Moving From Awareness to Action?

The disconnect between students' awareness about civic issues and their action towards them is known as the "commitment" gap. This is because many students understand the "why" of an issue but fail to understand the actions that are required to make change. Some students feel powerless: believing their individual actions are ineffective. Limited participation by students can also be caused by a high cost of sustainable options along with a lack of time. These students do not receive guides in tackling these civic issues, limiting the amount of students who take real-world action (Benzehaf et al., 2025).

Does Online Activism Build Long-Term Civic Identity, or Does It Fade When Trends Change?

Whether online activism can build long-term civic identity or fade when trends change depends on the individual, and often this activism contains a bit of both. But while topics may change with the algorithm, the habit of being a civic actor sticks.

How Can Digital Campaigns Be Designed to Encourage Sustained Engagement Rather Than Temporary Attention?

Digital campaigns can be designed to encourage sustained engagement by providing clear steps and replacing abstract facts with practical "how-to" guides which can turn concern into civic habits.

VIII. Equity and Access: Who Gets to Participate?

Does Digital Civic Engagement Give More Voice to Youth Who Are Too Young to Vote?

Digital civic engagement significantly expands the voice of young people including those who are too young to vote by allowing them to participate in democratic processes in non-traditional ways. Digital platforms allow youth to mobilize and advocate for policy change, even if not in voting booths; this creates a space which promotes active civic literacy (Nature Index, n.d.).

How Does Online Activism Help Students From Marginalized Communities Share Experiences?

Online activism gives students from marginalized communities an opportunity to engage in psychopolitical resistance. Psychopolitical resistance is a type of civic engagement which centers personal well-being to challenge systemic oppression. Online activism helps reshape narratives which previously silenced their perspectives. Online activism builds spaces for dialogues about discrimination, microaggressions, etc. which can provide support to marginalized communities (Wilf & Wray-Lake, 2021).

What Digital Divides Still Exist in Terms of Internet Access, Device Access, Privacy, Language, and Media Literacy?

Digital divides continue to exist as a form of systemic inequality due to financial constraints, infrastructure disparities, and gaps in digital literacy which disproportionately affect marginalized communities. Rural and indigenous areas may not have the same reliable broadband infrastructures as others, often lacking fiber optic cables and cell towers. Urban areas also face limitations due to landlords' limiting ISP agreements, limiting competition and leaving individuals with poor service. Digital redlining is also a major divide where internet service providers discriminate in the maintenance of infrastructure depending on the socioeconomic status, race, or ethnicity of a neighborhood. Microsoft's digital equity study estimates that 120.4

million Americans do not use the internet at broadband speeds. Affordability remains a huge barrier as well due to the cost and management of devices, software subscriptions, along with data plans and internet access. There is also a divide in linguistically inclusive content; there is no universal obligation for online apps/websites to adopt accessibility and translation tools. The literacy gap also limits digital skills, preventing individuals from finding and interpreting information online (Hollimon et al., 2025).

Are Some Students Punished or Judged More Harshly for Online Civic Expression?

Some students are punished or judged more harshly for online civic expression as it allows speech to be closely monitored and regulated by authorities. School authorities are especially prone to contextual misunderstandings where expressions understood in person can be misinterpreted and lead to punishment. They can also experience a loss of opportunities due to judgement towards unpopular values (American Bar Association, n.d.).

How Do Race, Class, Gender, Immigration Background, and Geography Shape Youth Digital Participation?

Digital participation is shaped by factors like race, class, gender, and geography. Geography remains a significant barrier to participation; this means that those with limited access may not participate digitally at all, especially in civic spaces. Low-income youth are also more likely to share devices among family members which limits their time for civic engagement online (Magis-Weinberg et al., 2025).

Can Online Civic Spaces Create Belonging for Youth Who Feel Excluded From Traditional Institutions?

Online civic spaces can create belonging for youth who feel excluded from traditional institutions. They lower traditional barriers and create room for every individual's voice to be heard.

Who Benefits Most From Digital Activism, and Who Is Left Out?

However, not everyone benefits equally from digital activism. Those from higher income households with higher education access and access to the internet are more represented digitally than those without these resources.

IX. Mental Health, Burnout, and Civic Fatigue

How Does Constant Exposure to Social Problems Affect Young People Emotionally?

Constant exposure to social problems can affect youth emotionally by creating higher levels of mental distress and increasing anxiety. It introduces negative heartbreaking content an individual wouldn't typically see in daily life and can make situations seem bleak, negatively impacting mental health.

What Is "Compassion Fatigue," and How Might It Affect Youth Activists?

"Compassion fatigue" is a state of being emotionally/physically drained as a result of caring for other people. For youth activists who invest heavy emotional labor into civic causes, compassion fatigue can show up in different ways. Activists can experience a loss of empathy and begin to feel disillusionment with the movement. It can also lead to feelings of powerlessness and sadness which can lead to behavioral withdrawal as well as guilt (Youth Intervention Programs Association, 2024).

Are Young People Pressured to Publicly Respond to Every Major Issue?

Young people may feel pressured to publicly respond to major issues, especially if they already have a following on social media. Young influencers and content creators are often publicly pressured to pick a side on divisive, nuanced issues.

Can Silence Online Be Interpreted as Not Caring?

Digital silence is frequently interpreted as "not caring." The human brain is wired to seek closure and validation; when the recipient of digital silence fails to receive these things, they often "fill in the blanks" with negative assumptions. However, creators' silence online may not always be due to indifference, it may be used for self-preservation or boundary-setting (Mirza & Kumari, 2025).

How Can Youth Civic Engagement Be Encouraged Without Overwhelming Students?

Youth civic engagement can be encouraged without burnout by balancing action with student capability. School educators can focus on connecting assignments into real-world civic lessons to build a bridge without extra workload. Students can practice low-stress activism for local community issues. Youth can build communities of support to acknowledge small progress when big "wins" feel impossible. Online media can often make issues feel overbearing and solutions seem impractical; it is important for students to educate themselves in doses and take a break online if necessary.

X. Gaps in the Research

Too much research treats "social media" as one category instead of separating TikTok, Instagram, YouTube Shorts, Snapchat, and X. Platforms significantly differ in their digital architecture which determines how users interact with them. For example, while 90% of teens use YouTube, only 31% use Facebook and 24% use WhatsApp, yet a disproportionate amount of civic research remains focused on Facebook (Boulianne, 2022).

More research is needed on younger teens, not just college students or eligible voters. As children and adolescents enter social media at a critical development stage, it is important that more research is inclusive to these groups. Data shows that 12–27% of children discuss political content online, yet their specific motivations and barriers remain understudied (UNICEF Office of Research – Innocenti, 2020).

There is limited research on short-form video algorithms specifically. While YouTube now averages over 70 billion daily views and 63% of Gen Z engage with these platforms daily, most media literacy research remains focused on long-form traditional news (Yu et al., 2025).

More attention is needed on marginalized youth and students who cannot participate in traditional civic spaces. Digital civic engagement is more equitable than traditional forms; it gives most groups equal access among different racial and ethnic groups. For youth holding marginalized identities, digital spaces serve as a square for psychopolitical resistance, however they are also more vulnerable to digital surveillance, harassment, and "contextual misreading" by institutional authorities (Wilf & Wray-Lake, 2021).

Researchers need better ways to measure micro-activism beyond likes, shares, and comments. Failing to recognize expressive participation (such as posting original thoughts or asking critical questions) leads to an underestimation of youth political involvement. Current metrics also often miss civic acts that do not fit into categories like voting or party affiliation. Traditional measures are failing to capture how youth actually interact with the world (Sormanen, 2025).

XI. Conclusion

What Should Readers Understand Differently About Youth Civic Engagement After Reading This Paper?

After reading this paper, individuals should have a fundamentally different definition of youth civic engagement, specifically regarding its broadened definition, psychological impact, and the systemic forces that shape it.

Civic engagement is no longer just institutional. Digital networked participation is a legitimate and crucial form of civic engagement. Liking, reposting, or commenting on civic posts are forms of "micro-activism" which provide a low barrier entry point for youth who are too young to vote or face physical/financial barriers to traditional participation.

Modern youth often passively encounter civic information through recommendation algorithms. This creates a reciprocal loop where interactions feed similar content. This "chooses" the civic issues youth will see; this can help students discover topics they won't see in school, but it can also risk over-exposure to emotionally intense partisan content over nuanced explanations and perspectives.

Research shows that online and offline activism is deeply intertwined. Digital participation is a significant indicator of future offline behavior. Its impact is neither automatically shallow or automatically powerful; its impact depends on the connection and bridge between online attention and real-world opportunities.

How Can Society Help Young People Become Informed Digital Citizens Instead of Passive Content Consumers?

Society can help youth become informed digital citizens rather than passive content consumers by providing tools that bridge the commitment gap between caring about an issue and knowing how to take action. One such tool is an app called “Sway – Civic Engagement” intended to bridge this gap. It provides legislation tracking such as bill status and voting history, actionable scripts (including phone and email templates) to contact legislators, and helps educate users on the legislative process. Beyond using such platforms, educational institutions have a primary responsibility to teach civic media literacy. Current education is often criticized for avoiding controversial current topics and using outdated tools for web credibility. Schools should teach lateral reading, build political efficacy, and integrate real-world civic lessons into assignments. Youth organizations and non-profits should mobilize individuals and create communities of support. These changes can further bridge any disconnect between online and physical civic spaces.

Works Cited

Akash & Chowdhury. (2025). Small changes, big impact: A mini review of habit formation and behavioral change principles. https://www.researchgate.net/publication/391205475

American Bar Association. (n.d.). Online speech and its impact on students' ability to openly discuss public issues. Human Rights Magazine. https://www.americanbar.org/groups/crsj/resources/human-rights/archive/online-speech-impact-students-ability-openly-discuss-public-issues/

Benzehaf et al. (2025). Exploring university students' perceptions and engagement in environmental awareness and sustainable practices. Discover Sustainability. https://link.springer.com/article/10.1007/s44274-025-00462-w

Boulianne. (2022). Standby Ties that Mobilize: Social Media Platforms and Civic Engagement. https://pmc.ncbi.nlm.nih.gov/articles/PMC10240620/

Branje et al. (2021). Dynamics of Identity Development in Adolescence: A Decade in Review. https://pmc.ncbi.nlm.nih.gov/articles/PMC9298910/

Brodsky et al. (2021). Improving college students' fact-checking strategies through lateral reading instruction in a general education civics course. Cognitive Research: Principles and Implications. https://link.springer.com/article/10.1186/s41235-021-00291-4

Cooper. (2023). The Effectiveness of Online Activism: Who It Is Effective for, What Issues It Is Effective for, and What Time Period It Is Effective for [Honors thesis]. University at Albany, State University of New York. https://scholarsarchive.library.albany.edu/cgi/viewcontent.cgi?article=1042&context=honorscollege_pos

De et al. (2025). Social Media Algorithms and Teen Addiction: Neurophysiological Impact and Ethical Considerations. https://pmc.ncbi.nlm.nih.gov/articles/PMC11804976/

Galais & Anduiza. (n.d.). The slacktivism crossroad: causal relationships between online and offline political participation. Observatory on Participatory Democracy. https://www.oidp.net/docs/repo/doc38.pdf

Gentry. (2025). The psychology of viral content: Why we share. https://www.comgroup.com/blog/the-psychology-of-viral-content-why-we-share

Greijdanus (2020). The psychology of online activism and social movements: relations between online and offline collective action. Current Opinion in Psychology. https://www.sciencedirect.com/science/article/pii/S2352250X20300324

Hauta et.al. (2021). Showing they don't care [PDF]. Michigan State University. https://hrcc.cas.msu.edu/news/showing-they-dont-care.pdf

Hollimon et al. (2025). Redefining and solving the digital divide and exclusion to improve healthcare: going beyond access to include availability, adequacy, acceptability, and affordability. https://pmc.ncbi.nlm.nih.gov/articles/PMC12052546/

Hughes. (n.d.). Small Acts, Bigger Impact: What Is Microactivism?. https://www.proudchangemakers.org/voice/ocean-e853m

Klüser. (2025). From Entertainment to Engagement? Entertainment Figures' Political Messaging and Audience Responses in the Digital Age. Political Communication. https://www.tandfonline.com/doi/full/10.1080/10584609.2025.2498526

Knifton et al. (2024). Young People's Interest in Voting and Civic Engagement: Interviews with Adolescents in the Midwest. University of Chicago. https://bpb-us-w2.wpmucdn.com/voices.uchicago.edu/dist/1/4039/files/2025/01/Voting-Brief_6.21.24.pdf

Kornbluh et al. (2022). Exploring civic behaviors amongst college students in a year of national unrest. https://pmc.ncbi.nlm.nih.gov/articles/PMC9545712/

Laursen & Veenstra. (2021). Toward understanding the functions of peer influence: A summary and synthesis of recent empirical research. https://pmc.ncbi.nlm.nih.gov/articles/PMC8630732/

Lee & Xenos. (2022). Incidental news exposure via social media and political participation: Evidence of reciprocal effects. https://www.researchgate.net/publication/346097621

Magis-Weinberg et al. (2025). Global Perspectives on Youth and the Digital Environment: Learnings from Majority World Countries. In: Christakis, D.A., Hale, L. (eds) Handbook of Children and Screens. Springer. https://link.springer.com/chapter/10.1007/978-3-031-69362-5_69

Metzler & Garcia. (2024).Social Drivers and Algorithmic Mechanisms on Digital Media. https://pmc.ncbi.nlm.nih.gov/articles/PMC11373151/

Milli et al. (2025). Engagement, user satisfaction, and the amplification of divisive content on social media. https://pmc.ncbi.nlm.nih.gov/articles/PMC11894805/

Mindel et al. (2024). Digital activism to achieve meaningful institutional change: A bricolage of crowdsourcing, social media, and data analytics. Research Policy. https://www.sciencedirect.com/science/article/abs/pii/S0048733323002354

Mirza & Kumari. (2025). The Power of Digital Silence. Zenodo. https://zenodo.org/records/15779613

Muringa & Adjin-Tettey. (2024). Media Literacy's Role in Democratic Engagement and Societal Transformation among University Students. African Journalism Studies. https://www.tandfonline.com/doi/full/10.1080/23743670.2024.2424902

Nature Index. (n.d.). Digital citizenship and civic engagement among youth. Nature. https://www.nature.com/nature-index/topics/l4/digital-citizenship-and-civic-engagement-among-youth

Oinonen & Paloniemi. (2023). Understanding and measuring young people's sustainability actions. Journal of Environmental Psychology. https://www.sciencedirect.com/science/article/pii/S027249442300172X

Olaniran & Williams. (2020). Social Media Effects: Hijacking Democracy and Civility in Civic Engagement. https://pmc.ncbi.nlm.nih.gov/articles/PMC7343248/

Park et al. (2023). Exploring the relationship between media literacy, online interaction, and civic engagement. The Information Society. https://www.tandfonline.com/doi/full/10.1080/01972243.2023.2211055

Pew Research Center. (2025). Teens, social media and AI chatbots: 2025. https://www.pewresearch.org/internet/2025/12/09/teens-social-media-and-ai-chatbots-2025/

Pew Research Center. (2025). Young adults and the future of news. https://www.pewresearch.org/journalism/2025/12/03/young-adults-and-the-future-of-news/

Rachimoellah et al. (2024). Digital Activism and Political Change: Challenges of Social Media's Impact on Political Development. Media and Environmental Information Studies. https://scholarhub.ui.ac.id/cgi/viewcontent.cgi?article=1107&context=meis

Schmitt et al. (2024). What does civic education do in a digitalized world?. https://pmc.ncbi.nlm.nih.gov/articles/PMC10961436/

Sormanen. (2025). Young people's online political participation in an apolitical digital sphere: identifying political topics and expression patterns on a youth discussion forum. Journal of Youth Studies. https://www.tandfonline.com/doi/full/10.1080/13676261.2025.2581983

Stahl. (2025). Social media influencers blur the lines between political content and campaigning, potentially affecting elections. Yahoo News. https://ca.news.yahoo.com/social-media-influencers-blur-lines-121915956.html

Stahl & Hu. (2025). Authenticity, activism, and the attention economy: Deciphering how Australian young people navigate the social pressures of a techno-social civic landscape. International Journal of Educational Research. https://www.sciencedirect.com/science/article/pii/S088303552500223X

Stanford Graduate School of Education. (2016). Stanford researchers find students have trouble judging the credibility of information online. https://ed.stanford.edu/news/stanford-researchers-find-students-have-trouble-judging-credibility-information-online

Swart. (2021). Experiencing Algorithms: How Young People Understand, Feel About, and Engage With Algorithmic News Selection on Social Media. Social Media + Society. https://journals.sagepub.com/doi/10.1177/20563051211008828

UNICEF Office of Research – Innocenti. (2020). Digital civic engagement by young people. https://www.unicef.org/innocenti/reports/digital-civic-engagement-young-people

UTHealth Houston School of Public Health. (2025). New study finds viewing political TikTok influencers can increase negative emotional affect in young adults. https://sph.uth.edu/news/story/new-study-finds-viewing-political-tiktok-influencers-can-increase-negative-emotional-affect-in-young-adults

Wang and Sundar. (2022). Liking versus commenting on online news: effects of expression affordances on political attitudes. Journal of Computer-Mediated Communication, 27(6), zmac018. https://academic.oup.com/jcmc/article/27/6/zmac018/6700671

Wilf & Wray-Lake. (2021). "That's How Revolutions Happen": Psychopolitical Resistance in Youth's Online Civic Engagement. https://pmc.ncbi.nlm.nih.gov/articles/PMC12443231/ (see also https://pubmed.ncbi.nlm.nih.gov/40969284/ — confirm whether these are the same record)

Youth Intervention Programs Association. (2024). Compassion fatigue in youth work and ways to manage it. https://training.yipa.org/quick-takes-blog/compassion-fatigue-in-youth-work-and-ways-to-manage-it/

Yu et al. (2025). Enhancing media literacy to combat information fragmentation in digital short video platforms: a cross-sectional study. https://pmc.ncbi.nlm.nih.gov/articles/PMC12770493/

Artificial Intelligence Personalization and the Changing Social Dynamics of Sports Fandom
By Simon Zhong

Abstract

Artificial intelligence is changing the way sports organizations connect with fans. Teams can now use fan data and AI tools to decide what content different fans see and how they interact with sports online. This paper looks at what those changes may mean for sports fandom from a sociological perspective, with a particular focus on the NFL. Using a narrative review of research from sports sociology, marketing, digital media, consumer behavior, and AI ethics, the paper focuses on three areas: fan identity and community, engagement and loyalty, and trust. The research shows that personalization can make fans feel more connected to the teams and players they already follow, but it can also make each fan's experience more personal and different. Algorithms may lead to more fan engagement, but more engagement does not always mean stronger fan loyalty. Questions about privacy, transparency, and the use of fan data may also affect how comfortable fans are with these technologies. AI personalization may be changing the way fans experience being part of the same team community, with fans still sharing the same team identity but experiencing it in different ways.

Keywords artificial intelligence, AI personalization, sports fandom, fan identity, fan engagement, fan loyalty, algorithmic curation, digital communities

Introduction

Background

Artificial intelligence is changing how professional sports organizations communicate with fans. Teams can use fan data, recommendation systems, predictive analytics, and machine learning to provide content, merchandise suggestions, and digital experiences that match individual interests (Rathinavel 3; Westerbeek). These tools can help fans stay connected to their teams even when there is no game going on. But as the experience becomes more personalized, it may also change how fans connect with each other and experience being part of a fan community. Fandom has traditionally involved shared experiences, such as watching games together, following the same team stories, and celebrating or losing as a group. As more of the fan experience becomes personalized, different fans may begin to experience the same team in different ways. This makes us question how AI could change both fan identity and the sense of community they share with other fans.

Problem Statement

AI-driven personalization has become an important part of modern sports marketing, but most research has focused on how it can increase engagement, improve fan experiences, and help sports organizations perform better (Westerbeek; Rathinavel 3). Less research has looked at how

these technologies may change the social side of sports fandom. As recommendation systems, targeted ads, and personalized content become more advanced, different fans may start to experience the same team in different ways. These technologies may affect how fans develop their identities and experience collective belonging, form preferences and patterns of engagement, maintain loyalty, and judge whether increasingly personalized fan experiences are trustworthy.

This gap is particularly relevant to the NFL, where fandom, like sports fandom more broadly, has historically been sustained through collective experiences and shared emotional investment (Wann and Branscombe 1–2; Tajfel and Turner 15). Algorithmically curated content has the potential to strengthen engagement while simultaneously fragmenting shared experiences and narrowing exposure to diverse perspectives (Bozdag 209; Westerbeek). Broader algorithmic and sports-AI scholarship also raises concerns regarding privacy, surveillance, and the commercialization of fan data (Mittelstadt et al. 10; Westerbeek). Because of this, it is important to understand how AI-driven personalization may be changing the social side of sports fandom.

Purpose and Scope

This study looks at how AI-driven personalization may affect fan identity and belonging, engagement and loyalty, and trust and ethical concerns in sports fandom, with a particular focus on the NFL. Instead of looking at personalization mainly as a marketing strategy, the study focuses on how algorithms may change the social relationships and shared experiences that have traditionally been an important part of being a sports fan. Using ideas from sociology, sports marketing, and consumer behavior, this paper reviews existing research to better understand how AI personalization may be changing the way fans experience and take part in fandom. While the NFL is the main focus, this study also looks at how AI and digital platforms may be changing the way people connect and interact in society.

Research Questions

1. How does AI-driven personalization influence fan identification and the relationship between individual and collective experiences within sports fandom?
2. How does AI-driven content personalization shape fan engagement, preferences, and loyalty within increasingly individualized digital environments?
3. How do concerns about privacy, transparency, surveillance, and commercialization affect the trustworthiness and social legitimacy of AI-driven personalization in sports fandom?

Methodology

This study uses a narrative literature review to look at the sociological effects of AI-driven personalization in sports fandom, with a focus on the NFL. Instead of collecting data directly from fans, the study uses existing research from related fields, including sports sociology, sports marketing, consumer behavior, digital media, and artificial intelligence ethics.

This approach helps connect newer research on artificial intelligence with established research on sports fandom.

Google Scholar served as the main research tool for identifying relevant academic literature. Searches used combinations of terms related to the major themes of the study, including *AI personalization in sports*, *artificial intelligence and sports marketing*, *AI fan engagement*, *sports fan identification*, *sports fandom and social identity*, *algorithmic personalization and sports fans*, *fan loyalty*, *sports fan communities*, *recommendation systems in sports*, *fan data privacy*, and *AI ethics in sports*. Additional searches looked at related topics such as algorithmic curation, online fan communities, privacy, transparency, and trust. Sources were mainly chosen based on how well they connected to the study's main topics: fan identification and community, engagement and loyalty, and trust and ethical legitimacy.

The review uses both earlier and more recent research. Earlier studies on social identity and fan identification help explain the shared and social side of sports fandom. More recent research looks at digital engagement, algorithmic personalization, recommendation systems, and data ethics to understand how these relationships may be changing as fan experiences become more personalized. Because there is still limited research specifically on AI personalization in NFL fandom, this study also uses relevant research from professional sports more broadly, digital fandom, marketing, and algorithmic media.

Because this is a narrative review rather than a systematic review, the study does not try to cover every piece of research on artificial intelligence and sports fandom. Instead, it brings together findings from related areas to identify repeated patterns, tensions, and gaps in the existing research. The analysis follows the study's research questions and looks at how personalized digital experiences may affect the shared identity and sense of belonging that have long been part of sports fandom.

Review of Literature

AI-Driven Personalization and the Transformation of Sports Fandom

Recent advances in artificial intelligence have changed sports marketing by making fan engagement more personalized. Research on AI-enabled personalization shows that algorithmic systems can use customer information to customize interactions and recommendations across digital platforms (Gao and Liu 665–66). In professional sports, teams are using machine learning, predictive analytics, recommendation systems, and behavioral data to create more personalized fan experiences (Rathinavel 3; Westerbeek). These experiences can include customized highlights, targeted notifications, merchandise recommendations, and digital content based on a fan's past behavior and interests. Because of this, AI-driven personalization has become a more important part of how professional sports organizations connect with fans.

In professional sports, digital technology has expanded fan engagement beyond traditional marketing and game-day experiences (Vale and Fernandes 38–39). AI-driven personalization builds on this by adjusting interactions based on each fan's interests and

behavior. Westerbeek explains that personalized technologies can help sports organizations stay connected with fans by giving them relevant content and more chances to engage outside of major events. This suggests that AI personalization is more than just a marketing tool. It is also becoming part of a more continuous and personalized digital fan experience.

Existing research shows that AI-driven personalization can increase engagement and improve digital fan experiences, but less research has looked at its broader sociological implications (Rathinavel 3; Westerbeek). As algorithmic systems play a bigger role in shaping the content and interactions supporters encounter, personalization may affect not only how fans consume sports but also how they develop identification, participate in fan communities, form preferences, and maintain relationships with teams. Because of this, it is important to examine AI personalization through a broader sociological lens, especially its effects on fan identity and collective belonging, engagement and loyalty, and trust and ethical legitimacy.

Fan Identification, Collective Belonging, and Personalized Communities

Fan identification refers to the degree to which individuals incorporate a sports team into their self-concept, making team support an important component of both personal and social identity. Social Identity Theory helps explain this by showing that people often form part of their identity through the social groups they belong to and care about (Tajfel and Turner 16). Within sports fandom, these group affiliations are reinforced through shared traditions, community participation, and collective emotional experiences. Wann and Branscombe found that fans who strongly identify with a team usually feel more emotionally connected to it and take part in more team-related activities (1). Taken together, this literature suggests that fan identity is both individual and collective: supporters personally incorporate a team into their sense of self while developing that identity through participation in a broader community of fans.

Recent research suggests that personalized digital environments are giving fans more ways to strengthen fan identification. Research on AI-enabled personalization shows that behavioral information can be used to shape digital interactions around individual interests and patterns of fan engagement (Gao and Liu 665–66). In sports, personalized highlights, player-specific content, targeted notifications, and other individualized interactions may strengthen interests fans already have by keeping them connected to the teams and players they follow (Rathinavel 3; Westerbeek). Research on digital fan engagement also shows that relevant and easy-to-access online interactions can encourage fans to keep participating in sports communities (Vale and Fernandes 49). However, AI-mediated content does not always create a positive emotional response. Rebge and Nevzat, in their research with Generation Z football supporters, found that fans could see the usefulness of AI-mediated sports content but responded less positively when the content did not feel authentic or emotionally realistic. Overall, personalized digital experiences may help strengthen fan relationships, but their emotional impact may depend on whether the content feels relevant and authentic to fans.

The literature shows that AI-driven personalization is adding new ways for fan identification to develop rather than replacing the traditional ones. Fan identity may now develop

through both collective experiences and personalized digital interactions, allowing fans to stay connected to the same team while experiencing that connection in different ways. This raises a larger question about how personalization may be changing the fan communities that have traditionally been part of sports fandom.

Sports fandom is not only about supporting a team; it is also a social experience built through shared traditions, common stories, and being part of a larger fan community. Fans can develop a sense of belonging by going to games, talking about team performance, taking part in traditions, and sharing wins and losses with other supporters. These shared experiences help build a collective identity across the fan base. As personalization becomes a bigger part of digital fandom, the question is whether more individualized fan experiences will strengthen this sense of belonging or begin to change the shared experiences that sports fandom has traditionally depended on.

Digital platforms have made it easier for fans in different places to stay connected to teams and fan communities throughout the season (Vale and Fernandes 39). AI-driven personalization builds on this by shaping content and interactions around individual interests and patterns of fan engagement. Westerbeek suggests that these personalized experiences can give fans more frequent and relevant ways to interact with sports organizations, while research on social-media engagement shows that digital communication can also encourage participation in online fan communities (Ichsan et al. 458). Together, these changes have made sports communities more accessible by reducing geographical barriers and giving fans more ways to take part in sports fandom online.

At the same time, personalized digital environments may be changing the shared side of sports communities. Because recommendation systems use each fan's interests and past behavior, different fans may start seeing different sports content. Westerbeek warns that algorithmic recommendation systems can strengthen digital echo chambers and limit exposure to other sports content. These personalized experiences can make sports fandom feel more relevant to individual fans, but they may also reduce some of the shared stories and experiences that have traditionally connected the larger fan community. At the same time, personalization can bring more fans into sports fandom while also creating smaller groups around certain interests, players, or ways of engaging. Sports fandom may start to look more like smaller connected fan communities that support the same team, even though fans may experience that support in different ways.

Algorithmic Curation, Fan Engagement, Preferences, and Loyalty

AI-driven personalization can change sports fandom by affecting what different fans see online. Recommendation systems may show different content to different fans. For example, they may show different highlights, news stories, player profiles, discussions, and ads based on what a fan has looked at or followed before (Westerbeek). This means that two people who support the same team may still see very different content about that team. Certain players, stories, or topics may appear more often for one fan than for another. Algorithms can shape what fans see, but fans still decide what that content means to them and how they respond to it.

Because personalized content affects what fans see over and over, it may also shape their interests and how they engage over time. Broader digital-media scholarship has raised concerns that personalization can narrow exposure to alternative information by prioritizing content consistent with previous interests (Bozdag 209). Sports-specific scholarship similarly warns that algorithmic recommendation systems may reinforce digital echo chambers and reduce exposure to alternative sporting content (Westerbeek). Within sports fandom, these processes may influence what supporters choose to watch, follow, and discuss by repeatedly prioritizing content that aligns with previous interests. Recommendation systems do not take away fan choice, but they can shape the information fans see when they form preferences and make decisions. By repeatedly showing certain types of content, algorithms may influence fan engagement without directly controlling what fans choose.

These patterns of fan engagement are also important for understanding fan loyalty in more personalized digital environments. Traditionally, fan loyalty has meant a strong and lasting commitment to a team based on emotional attachment, being part of a fan community, and identifying with the team over time (Wann and Branscombe 1–2). More recent research suggests that AI-driven personalization may strengthen behaviors connected to loyalty by encouraging fans to interact more often through personalized content, targeted communication, and individualized digital experiences (Rathinavel 3; Westerbeek). However, interacting more often does not always mean that fans feel a deeper emotional commitment. Rebge and Nevzat found something similar in their research with Generation Z football supporters: fans could see AI-mediated sports content as useful without necessarily having the same positive emotional response or choosing to use it more. Personalization may lead fans to interact with teams and related content more often, but there is still less evidence showing that this creates stronger long-term identification with a team. Because of this, AI-driven personalization seems more clearly connected to behavioral engagement and behavioral forms of fan loyalty than to the deeper commitment traditionally connected with collective sports fandom.

The literature suggests that AI-driven personalization does not directly determine what fans choose, but it can shape the environment where their preferences and patterns of engagement develop. Through algorithmic curation, fans may keep seeing content that matches their existing interests, which can lead to more repeated interaction and stronger behavioral forms of loyalty. However, it is still unclear whether this leads to deeper, long-term commitment. Fans still make their own choices, but algorithms can influence which parts of sports fandom get more of their attention. As personalized systems become a bigger part of the fan experience, this also raises questions about whether fans see these systems, and the organizations using them, as transparent, responsible, and trustworthy.

Privacy, Transparency, Trust, and Ethical Legitimacy

The growing use of AI-driven personalization creates a tradeoff between making fan experiences more relevant and collecting the large amounts of data needed to do that. Personalized systems may use things like viewing habits, purchase history, browsing behavior,

and other online activity to decide what content and recommendations different fans see. These practices can make digital engagement more useful and convenient, but they also raise concerns about privacy, surveillance, consent, and the commercialization of fan data (Westerbeek). Because of this, the ethical legitimacy of personalization depends not only on what these systems can offer, but also on how sports organizations collect, manage, and use their personal information.

Research on the ethics of algorithms often focuses on transparency and accountability. Mittelstadt et al. point to concerns about fairness, accountability, and transparency, especially when people do not fully understand how automated systems shape the information and choices they are given (6–10). Martin also argues that organizations are still responsible for the algorithms they create and use (844). Papagiannidis et al., based on a comparison of how organizations use AI, show the importance of having strong governance practices to build reliable AI systems and manage the risks that come with using them (123). In sports personalization, this means organizations should clearly explain how fan data is collected and used, as well as how personalized systems affect the content and recommendations fans receive. Being more transparent may help make AI-driven fan experiences feel more trustworthy and legitimate by giving fans a better understanding of the systems behind them.

The literature suggests that trust and ethical legitimacy will be important as AI-driven personalization becomes a bigger part of sports fandom. Personalized technologies can strengthen engagement and make fan experiences feel more relevant, but concerns about privacy, surveillance, transparency, and the commercial use of personal data may make fans less comfortable with these systems. Responsible data practices and greater transparency can help maintain trust between sports organizations and their supporters. As AI-driven personalization becomes more common in sports fandom, its future impact may depend on both how well the technology works and whether fans feel it is being used responsibly and with their interests in mind.

Discussion

Integrated Summary of Findings

The findings show that AI-driven personalization is changing sports fandom in ways that are more complicated than simply strengthening or weakening fan engagement. Instead, it seems to be changing the relationship between individual fan experiences and the collective identities and communities that have traditionally been part of sports fandom. Fans may still support the same teams and be part of the same larger fan communities, but they may experience those connections through different content, interests, and patterns of digital participation (Westerbeek). This may be creating a more hybrid form of fandom, where individualized digital experiences exist alongside collective belonging. These changes also go beyond identity and community, affecting how preferences are reinforced, how fan loyalty develops, and how much trust fans place in the systems shaping their experiences.

Hybrid Fandom: Individualization and Collective Belonging

AI-driven personalization seems to be changing the connection between individual and collective belonging, rather than taking one away from the other. Fans can still strongly identify with the same team and be part of a shared community, while different digital experiences can strengthen that connection in different ways. One fan may see personalized content about certain players, another may focus more on statistics or highlights, while another may connect through online discussions or other digital content. These different experiences do not necessarily make the larger group identity weaker. They show that collective fandom can still be important, even if fans experience and express their fan identity in different ways.

This hybrid form of fandom shows that community growth and fragmentation can happen at the same time. Personalized digital environments can make sports fandom more accessible and relevant to individual fans, while also reducing some of the shared stories and experiences across a larger fan base (Vale and Fernandes 39; Westerbeek). The result may be a fan community that is less uniform but still meaningful, with fans connected by the same team even if they engage through different interests, platforms, and forms of fan engagement. Collective belonging may depend less on fans having the exact same experiences but more on whether personalized experiences can still make them feel like part of the same group. In this way, AI-driven personalization may not replace traditional sports fandom, but it could change how collective identity is built and maintained over time.

Algorithmic Mediation, Engagement, and Loyalty

One important effect of AI-driven personalization may be that it increases repeated fan engagement without necessarily leading to deeper fan loyalty. Through algorithmic curation, content can become more relevant to individual fans and encourage them to return to certain teams, players, products, or types of content (Rathinavel 3; Westerbeek). These behaviors may look like loyalty, but frequent interaction does not always mean a fan has a stronger emotional commitment to a team. A fan may watch more personalized content or interact with a team more often without becoming more committed over time. This suggests that AI personalization may be more effective at increasing behavioral engagement than building deeper, long-term fan loyalty. Recent research with football supporters also found that seeing AI-mediated sports content as useful did not always lead to positive emotional responses (Rebge and Nevzat).

Personalization may affect what fans do, but it does not control their choices. Fans still make their own decisions and form their own preferences, even though recommendation systems help shape the digital content they see. When certain players, storylines, products, or types of content show up more often, fans may pay more attention to them, making them a bigger part of how they experience a team (Westerbeek). Over time, repeatedly seeing the same kinds of content may strengthen interests fans already have while making other content less important. The main point is that personalization does not take away fan choice, but it may influence how fan preferences, fan engagement, and fan loyalty develop over time.

Trust and the Sustainability of Personalized Fandom

Trust becomes more important as personalization changes the relationship between fans and sports organizations. In a traditional fan relationship, fans mainly connect with teams, players, and share sports experiences. AI-driven personalization adds another layer because sports organizations can collect behavioral data, predict what individual fans may prefer, and use those predictions to shape future interactions (Westerbeek). This means fans are not only connecting with a team, but also with the data systems working behind those interactions. Privacy and transparency matter because they can affect whether fans see personalization as helpful to their sports fandom or as something that feels too intrusive or commercial.

The long-term success of personalized fandom may depend on how well sports organizations balance relevance with accountability. More advanced technology does not automatically mean that personalized experiences will create stronger long-term relationships with fans. As organizations become better at predicting interests and personalizing interactions, how they use those abilities becomes more important. Personalization that feels useful and respectful of fan interests may strengthen the relationship between fans and sports organizations, while practices that feel intrusive or mainly commercial may weaken it. Trust is therefore more than just a reaction to personalization. It may also affect whether increased fan engagement develops into a lasting relationship between fans and sports organizations (Mittelstadt et al. 7; Martin 844; Papagiannidis et al. 125).

Contribution to Existing Research

This study adds to existing research by connecting ideas about fan identification, community, algorithmic personalization, loyalty, and AI ethics in one sociological framework. Instead of only looking at whether AI-driven personalization can increase engagement or improve marketing, this study also looks at how it may change the connection between individual fan experiences and collective belonging. The idea of hybrid fandom helps explain this change. Fans can still share a strong team identity and be part of larger fan communities, even if they get that connection in different ways. Instead of only asking whether personalization makes fandom stronger or weaker, this study looks at how it may change the way fan identity and community develop over time.

This study also looks at these changes as connected rather than separate. Personalized information environments may shape which parts of sports fandom get more attention, which can influence fan engagement and strengthen certain interests over time. These patterns may lead to behavioral forms of fan loyalty, while more individualized fan experiences may also change how collective identity and community are maintained. At the same time, these relationships may depend on whether fans see the systems behind personalization as trustworthy and legitimate. AI-driven personalization may therefore affect sports fandom in several connected ways involving fan identity, fan engagement, fan loyalty, and trust. Looking at these changes together can help explain how artificial intelligence may be changing sports fandom.

Limitations

Several limitations define the scope of this study. First, the study uses existing research instead of collecting primary data directly from sports fans. Because of this, the study cannot directly tell how individual NFL fans experience AI-driven personalization or whether more digital engagement actually leads to deeper, long-term fan loyalty. Second, this study uses a narrative rather than a systematic literature review. Although sources were selected for their relevance to the research questions, the review does not attempt to identify or evaluate every study related to AI personalization and sports fandom. Third, research specifically examining AI-driven personalization within NFL fandom remains limited, requiring the study to draw from broader scholarship in sports marketing, digital fandom, social media, consumer behavior, and AI ethics. These sources give useful ideas and evidence, but research from broader areas may not always apply directly to NFL fans. Finally, the rapid development of artificial intelligence limits the availability of long-term evidence and makes it difficult to determine whether current patterns represent lasting changes in fandom or short-term responses to emerging technologies. These limitations do not take away from the value of the patterns found in this review, but they show why more direct, long-term, and NFL-specific research is needed.

Future Research Directions

Future research should get more direct input from sports fans and look at how AI-driven personalization may change fandom over time. Surveys, interviews, and focus groups with NFL fans could give a clearer picture of how fans experience personalized content and whether it affects identification, community belonging, engagement, loyalty, and trust. Long-term studies could also show whether repeated personalized engagement leads to stronger emotional commitment or mostly just short-term activity.

Additional research could look at how different types of personalization affect fans in different ways. For example, researchers could compare personalization focused on teams with personalization focused on individual players. They could also look at whether algorithmic curation changes the stories and perspectives fans see, and how personalized digital experiences fit with traditional parts of sports fandom, like going to games, watching with friends, and taking part in community traditions. Research across different professional sports leagues and cultures could also help show which patterns are unique to NFL fandom and which reflect broader changes in digital sports communities.

Future research should also look at how transparency, privacy, and the use of fan data for business purposes affect trust in personalized systems. As AI becomes a bigger part of sports platforms, research with fans could help show when personalization feels useful and relevant and when it starts to feel too intrusive or commercial. Together, this research could provide stronger evidence for understanding how AI-driven personalization is changing the relationship between individual engagement and collective sports fandom.

Conclusion

This study looked at how AI-driven personalization may be changing sports fandom through its effects on fan identity and collective belonging, engagement and loyalty, and trust in more data-driven fan experiences. The research suggests that personalization does not simply make fandom stronger or weaker. Instead, it may be creating a more hybrid form of fandom, where fans are still connected to the same teams and communities but experience and express those connections through different personalized digital experiences. Algorithmic personalization seems to have a stronger connection to repeated engagement and behavioral forms of loyalty than to deeper, long-term emotional commitment. At the same time, concerns about privacy, transparency, surveillance, and commercialization may affect whether fans see these personalized experiences as trustworthy and legitimate. Overall, AI-driven personalization may be changing how fan identity, community, engagement, and trust develop within sports fandom. As these technologies become a bigger part of sports, understanding how personalized digital experiences affect collective belonging will be important for understanding the future of sports fandom.

Works Cited

Bozdag, Engin. “Bias in Algorithmic Filtering and Personalization.” Ethics and Information Technology, vol. 15, no. 3, 2013, pp. 209–227. https://doi.org/10.1007/s10676-013-9321-6.

Gao, Youjiang, and Hongfei Liu. “Artificial Intelligence-Enabled Personalization in Interactive Marketing: A Customer Journey Perspective.” Journal of Research in Interactive Marketing, vol. 17, no. 5, 2023, pp. 663–680. https://doi.org/10.1108/JRIM-01-2022-0023.

Ichsan, Muhamad Chaerul, et al. “Social Media and Sports: Engagement for Sports Fans.” Journal of Physical Education and Sports, vol. 11, no. 4, 2022, pp. 456–463. https://doi.org/10.15294/jpes.v11i4.70778.

Martin, Kirsten. “Ethical Implications and Accountability of Algorithms.” Journal of Business Ethics, vol. 160, no. 4, 2019, pp. 835–850. https://doi.org/10.1007/s10551-018-3921-3.

Mittelstadt, Brent Daniel, et al. “The Ethics of Algorithms: Mapping the Debate.” Big Data & Society, vol. 3, no. 2, 2016, pp. 1–21. https://doi.org/10.1177/2053951716679679.

Papagiannidis, Emmanouil, et al. “Toward AI Governance: Identifying Best Practices and Potential Barriers and Outcomes.” Information Systems Frontiers, vol. 25, no. 1, 2023, pp. 123–141. https://doi.org/10.1007/s10796-022-10251-y.

Rathinavel, Akilan. “The Impact of Generative Artificial Intelligence on Personalized Fan Engagement and Digital Content.” Journal of Sport Business and Management, vol. 1, no. 1, 2024, pp. 1–5. https://iaeme.com/Home/article_id/JSBM_01_01_001.

Rebge, Suleyman, and Raziye Nevzat. “Between Passion and Prediction: AI, Identity, and Fan Engagement among Generation Z Football Supporters.” Sport Marketing Quarterly, 6 Apr. 2026, https://doi.org/10.1177/10616934261433377.

Tajfel, Henri, and John C. Turner. “The Social Identity Theory of Intergroup Behavior.” Psychology of Intergroup Relations, edited by Stephen Worchel and William G. Austin, 2nd ed., Nelson-Hall, 1986, pp. 7–24. https://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Intergroup_Conflict/Tajfel_%26_Turner_Psych_of_Intergroup_Relations_CH1_Social_Identity_Theory.pdf.

Vale, Leonor, and Teresa Fernandes. “Social Media and Sports: Driving Fan Engagement with Football Clubs on Facebook.” Journal of Strategic Marketing, vol. 26, no. 1, 2018, pp. 37–55. https://doi.org/10.1080/0965254X.2017.1359655.

Wann, Daniel L., and Nyla R. Branscombe. “Sports Fans: Measuring Degree of Identification with Their Team.” International Journal of Sport Psychology, vol. 24, no. 1, 1993, pp. 1–17. https://www.researchgate.net/publication/232572704_Sports_fans_Measuring_degree_of_identification_with_their_team.

Westerbeek, Hans. “Algorithmic Fandom: How Generative AI Is Reshaping Sports Marketing, Fan Engagement, and the Integrity of Sport.” Frontiers in Sports and Active Living, vol. 7, 2025, article 1597444. https://doi.org/10.3389/fspor.2025.1597444.

How Does Patent and Copyright Law Affect Innovation in the Video Games Industry
By Almira Seissembayeva

Abstract

This study examines how patent and copyright law affect innovation in the video game industry. The research uses a mixed-methods approach combining a review of academic and legal literature, comparative legal analysis, a semi-structured interview with a game developer, and an online survey of 225 individuals involved in game development. The findings suggest that intellectual property law has a significant but complex influence on game development. While 48.0% of survey respondents reported that copyright and patent law heavily or very heavily affected their work, only 34.7% agreed that current intellectual property laws benefit them directly. Furthermore, 48.4% of respondents believed that the current system benefits larger companies more than independent developers. The interview similarly indicated that larger companies have greater financial and legal resources to manage intellectual property issues. Overall, the findings suggest that intellectual property protection can encourage innovation by protecting creative and technological developments, but may also create practical barriers, particularly for independent developers. The study concludes that effective intellectual property regulation should balance the protection of creators with the need to maintain opportunities for experimentation and innovation.

Introduction

The video game industry has become one of the world's largest creative industries, generating billions of dollars annually and driving technological and artistic innovation. As games become increasingly complex, collaborative, and allow for more user interaction, intellectual property (IP) law plays a growing role in protecting creative works while also shaping the conditions under which future developers can create new content. Copyright and patent law are therefore not only legal mechanisms but also important factors that influence innovation within the industry.

Previous research has examined various aspects of intellectual property in video games, including copyright protection, patent disputes, user-generated content, and the legal ownership of digital assets. Scholars generally agree that intellectual property rights encourage investment by protecting creators' work. However, they also note that excessive legal protection may restrict creativity, increase legal uncertainty, and create barriers for smaller developers.

Using a mixed-methods approach that combines literature review, legal analysis, a semi-structured interview with a game developer, and an online survey of individuals involved in game development, the research seeks to evaluate how different intellectual property frameworks influence creative practices and innovation in the industry.

Literature Review

Lastowka (2013) examines how copyright law has evolved alongside the video game industry, focusing on the growing role of player-generated content. The author argues that

activities such as modding, fan art, and user-created game content challenge traditional copyright principles and require a more flexible legal approach. A major strength of the article is that it combines legal analysis with examples from the video game industry, showing how players have become creators rather than just consumers. This helps explain why copyright law is becoming increasingly important as games become more interactive. However, the study was published in 2013 and does not discuss more recent developments, such as modern intellectual property disputes or the rapid growth of live-service games. In addition, it focuses mainly on copyright law and does not compare different national legal systems. This article provides a useful theoretical foundation for our research by explaining the relationship between copyright law and creativity in video games.

Pengtang (2025) examines the relationship between copyright protection and innovation in the video game industry through the case of Black Myth: Wukong. The study focuses on the difficulty of protecting non-textual elements such as gameplay mechanics, visual styles, and interactive layouts under Chinese copyright law. The author argues that the existing fragmented approach can allow imitation while also creating uncertainty about the boundaries of copyright protection. A major strength of the study is its comparative legal perspective. The author examines Chinese copyright law alongside approaches in the United Kingdom and United States, particularly their treatment of the distinction between ideas and expressions. The study also connects legal protection to innovation, arguing that both underprotection and overprotection can create problems: weak protection may encourage imitation, while excessive protection may restrict further creativity. The article is particularly useful for the present research because it demonstrates how different approaches to copyright protection can affect the balance between protecting creators and maintaining opportunities for innovation. However, its primary focus is the Chinese legal system and piracy, rather than the experiences of game developers.

Goltz et al. (2015) examine the relationship between innovation, user satisfaction, and copyright in the video game industry by comparing *Minecraft* and *Call of Duty*. The authors focus on how user-generated content has changed the traditional relationship between game developers and consumers, creating new challenges for copyright law. A major strength of the study is its empirical approach. Rather than discussing copyright only theoretically, the authors analyze game maps and data from online communities to examine innovation and user satisfaction. This provides practical evidence about how user-generated content can influence the video game industry. However, the study has some limitations. It focuses on only two games and does not compare different national legal systems. Additionally, the researchers found no significant correlation between innovation and user satisfaction, meaning that further research is needed to understand how innovation affects market success. The article is useful for the present study because it demonstrates how copyright law and user-generated content can interact with innovation in the video game industry.

The existing literature suggests that intellectual property law has a dual relationship with innovation in the video game industry. Research on copyright and user-generated content demonstrates that legal protection can become increasingly difficult to apply as games

incorporate interactive and user-created elements, while Pengtang's analysis of *Black Myth: Wukong* shows that insufficient protection can facilitate imitation and potentially undermine incentives for innovation. At the same time, the comparison of different legal systems suggests that stronger protection is not necessarily preferable, as excessive protection may restrict the use of common ideas and creative foundations.

Taken together, these studies indicate that the central issue is not simply whether video game IP should be protected, but how the scope of that protection should be defined. However, existing research tends to focus on legal doctrine, copyright disputes, or consumer behavior rather than combining comparative legal analysis with the perspectives of developers themselves.

Methodology

This study employs a mixed-methods research design, combining qualitative and quantitative approaches to examine how patent and copyright law influence innovation in the video game industry.

The study relies on both secondary and primary data. Secondary data consist of academic journal articles, books, legal documents, court cases, and reports on intellectual property and the video game industry. These sources provide the theoretical and legal foundation for understanding how copyright and patent regulations shape innovation.

Primary data are collected through two research methods. First, a semi-structured interview is conducted with an experienced game developer currently involved in the development of the indie game *Carnival Hunt*. The interview explores the participant's experiences with copyright and patent law, the influence of intellectual property regulations on creative decision-making, and perspectives on innovation within the industry. A semi-structured format was chosen to ensure consistency across core questions while allowing the participant to elaborate on issues based on professional experience.

Second, an online questionnaire is distributed among individuals involved in game development, including professional developers, indie developers, programmers, and students studying game development. The survey consists primarily of multiple-choice questions and five-point Likert scale items designed to measure participants' awareness of intellectual property law, experiences with copyright and patents, and perceptions of how legal protection affects innovation.

The qualitative interview data are analyzed using thematic analysis, in which recurring ideas and themes are identified and interpreted. The quantitative survey data are analyzed using descriptive statistics, including frequencies and percentages, to identify overall patterns in participants' responses.

All participants are informed about the purpose of the research before participating. Participation is voluntary, responses to the questionnaire are anonymous, and interview participants provide informed consent prior to the interview. The collected data are used solely for academic research purposes.

Results & Discussion

Semi-structured interview

To complement the literature review, a semi-structured interview was conducted with an experienced game developer who has worked professionally in the industry for over five years and currently serves as a programmer on the indie horror game *Carnival Hunt*. The interview provided practical insights into how intellectual property (IP) law affects game development beyond the theoretical perspectives discussed in previous studies.

One of the most notable findings was that patents appear to play a relatively limited role in everyday game development. According to the interviewee, gameplay mechanics themselves are rarely patented because they are difficult to protect under existing legal frameworks. Instead, patents are more commonly associated with underlying technologies or hardware. This observation suggests that patents may influence technological innovation more than creative game design, supporting the view that legal protection in the gaming industry extends beyond gameplay mechanics alone.

The interview also revealed that copyright law has its greatest practical impact on the technical side of development. The participant described situations in which permissions were required to implement technologies such as NVIDIA DLSS, AMD FSR, and PlayStation controller support. These legal and administrative procedures were reported to slow development, particularly for independent studios that lack dedicated legal teams. This finding indicates that compliance with intellectual property regulations may create greater barriers for smaller developers than for large publishers.

Another important theme concerned open-source software. The participant reported frequent use of open-source tools and plugins in both commercial and independent projects and described them as an essential part of modern game development. Open-source technologies were viewed as reducing development costs, accelerating production, and encouraging collaboration within the developer community. These findings suggest that innovation is influenced not only by legal protection but also by the accessibility of shared technological resources.

Interestingly, the interview indicated that innovation is driven primarily by market competition rather than legal concerns. While intellectual property considerations were taken into account during the development of *Carnival Hunt*, the participant explained that the team's decision to differentiate its mechanics and artistic style from similar games was motivated mainly by the desire to establish a unique market identity. Avoiding copyright disputes was considered an additional benefit rather than the primary objective. This finding suggests that commercial incentives and creative ambitions may play a stronger role in encouraging innovation than the legal framework itself.

Finally, the participant emphasized that the practical burden of intellectual property law is not distributed equally across the industry. Large companies generally possess the financial and legal resources necessary to obtain licenses, negotiate agreements, and pursue litigation when necessary. Independent developers, however, often lack these resources and may therefore experience legal compliance as a greater obstacle. The interviewee also noted that smaller

creators rarely pursue legal action even when their own work is copied because litigation is often prohibitively expensive. This observation highlights an important gap between the formal legal protections provided by intellectual property law and their practical accessibility.

Overall, the interview supports the argument that intellectual property law has a complex and multidimensional relationship with innovation. Rather than simply encouraging or restricting creativity, its impact depends on the specific context. Copyright and patent regulations appear to protect technological innovation and creative works while simultaneously creating administrative and legal challenges, particularly for independent developers. These findings reinforce the view presented in the literature that intellectual property law must balance the protection of creators with the need to maintain an environment that allows innovation to flourish.

Survey

The survey collected 225 responses from individuals involved in the video game industry. The sample was predominantly composed of early-career developers: 140 respondents (62.2%) had 1–3 years of experience, while 65 (28.9%) had less than one year. Most participants worked in semi-large teams (56.4%), and programmers represented the largest professional group (46.2%).

Regarding familiarity with intellectual property law, 90 respondents (40.0%) described themselves as mostly or very familiar with copyright and patent law, while 49 (21.8%) reported being mostly or very unfamiliar. The largest group (38.2%) described their familiarity as moderate. This indicates that although many participants have some awareness of intellectual property law, high levels of legal familiarity are not universal among developers.

The survey showed that intellectual property law has a noticeable impact on participants' work. 48.0% of respondents stated that copyright and patents had heavily or very heavily affected their work, while 30.7% reported that they had barely affected it. However, perceptions of the benefits of the current system were more divided. Only 34.7% agreed or strongly agreed that current intellectual property laws benefit them, while 27.6% disagreed or strongly disagreed and 37.8% remained neutral.

A stronger pattern appeared when respondents were asked whether intellectual property laws benefit large companies more than independent developers. 48.4% agreed or strongly agreed with this statement, compared with 35.1% who disagreed or strongly disagreed. This perception is consistent with the interview, in which the developer noted that larger companies generally have greater financial and legal resources to manage licensing and intellectual property disputes.

The findings also complement previous literature, which has shown that the development of user-generated content and increasingly collaborative forms of game creation create new challenges for copyright law. Together, the survey and interview suggest that intellectual property law has a significant presence in game development, but its effects are not perceived uniformly. Rather than simply promoting or restricting innovation, the legal framework may provide protection while also creating different levels of difficulty for developers depending on their resources and position within the industry.

Research Limitations

The main limitation of the survey is the composition of the sample. More than 90% of respondents had fewer than four years of industry experience, while 46.2% were programmers and 56.4% primarily worked in semi-large teams. Therefore, the findings may not fully represent the perspectives of highly experienced developers, indie developers, or employees of large corporations. In addition, the survey measures participants' perceptions of intellectual property law rather than its objectively measured effect on innovation. The results should therefore be interpreted as evidence of industry perceptions rather than proof of a causal relationship between IP regulation and innovation.

Conclusion

This study examined how patent and copyright law affects innovation in the video game industry. The findings suggest that intellectual property law has a significant but complex influence on game development. Almost half of survey participants reported that copyright and patent law heavily or very heavily affected their work, yet only 34.7% agreed that the current system benefits them directly.

The findings also indicate a perceived imbalance between large companies and independent developers. Nearly half of respondents (48.4%) agreed that current intellectual property laws benefit larger companies more than indie developers. This perception was also reflected in the interview, where the developer emphasized the greater financial and legal resources available to large companies. However, the study does not establish that intellectual property law objectively disadvantages independent developers; rather, it demonstrates that such an imbalance is perceived by a substantial proportion of participants.

Overall, the research suggests that intellectual property protection should not be viewed as either inherently beneficial or harmful to innovation. Its effects depend on how legal protections operate in practice and on the resources available to those using them. A more effective intellectual property system should therefore protect creators and encourage investment while avoiding unnecessary barriers to experimentation, particularly for independent developers. Further comparative research involving more experienced developers and a more balanced sample could provide a clearer understanding of how the legal systems of different countries shape innovation in the industry.

Works Cited

Goltz, Nachshon and Franks, Jaimie and Goltz, Shem, Changing the (Video) Game: Innovation, User Satisfaction and Copyrights in Network Market Competition (March 30, 2015). Available at SSRN: https://ssrn.com/abstract=2587507 or http://dx.doi.org/10.2139/ssrn.2587507

Lastowka, Greg, Copyright Law and Video Games: A Brief History of an Interactive Medium (October 2013). Available at SSRN: https://ssrn.com/abstract=2321424 or http://dx.doi.org/10.2139/ssrn.2321424

Pengtang,X. (2025). A Study on the Economic Effects of Copyright Protection and Industrial Innovation in Black Myth: Wukong. Advances in Economics, Management and Political Sciences,191,171-178.

To What Extent Can Behaviours Commonly Associated with Psychopathy Be Explained by Social Roles, Socialization, and Institutional Environments? By Andy Peng

Abstract

Psychopathy is commonly understood as a biological or clinical condition, yet the role of social and environmental factors in shaping behaviours associated with it has received comparatively less attention in the research literature. This paper asks: to what extent can behaviours commonly associated with psychopathy — such as reduced empathy, moral disengagement, and manipulation — be explained by social roles, socialization, and institutional environments? Using a literature review methodology, this paper synthesizes findings across psychology, sociology, and criminology to evaluate three social explanations. Studies were selected based on their relevance to the three social explanations examined and their use of peer-reviewed empirical or theoretical evidence. Overall, the evidence suggests a consistent pattern, though the strength of that evidence varies across the three explanations. Social roles appear to bring out behaviours that resemble psychopathy in ordinary people fairly quickly, especially when someone holds a position of significant power with little accountability. Research has found associations between difficult childhood environments and higher levels of psychopathic traits, though research has not yet fully explained how that connection works, and institutional environments like prisons and the military appear to reinforce some of these patterns in adulthood, though much of this evidence is observational. Together, the evidence is most convincing when it comes to social environments drawing out or strengthening behaviours associated with psychopathy. The claim that they directly cause psychopathy itself is harder to support. What this means is that the way we design institutions, raise children, and structure organizations is not neutral. These choices may shape behaviour, and a clearer understanding of social influences could help inform efforts to reduce the conditions associated with these behaviours, though further research is needed to identify which approaches are most effective.

Introduction

When most people hear the word “psychopath,” they picture a dangerous individual, someone born “different” from the rest of us, born to be cruel and manipulative. Social culture has reinforced this image for decades, depicting psychopaths as monsters. The psychopath portrayed in society is someone who is an exception, someone whose life is so different that no amount of socialization, education, or even environmental change could have made them “normal.” However, this perspective raises a more challenging question. What if the behaviours we associate with psychopathy are not uniquely individual? What if social factors play a more significant role in producing these behaviours than we typically assume? This paper sets out to examine that question directly. Specifically, it asks, to what extent can behaviours commonly associated with psychopathy be explained by social roles, socialization, and institutional environments? The goal of this paper is not to reject biological explanations of psychopathy or to argue that psychopathic behaviour is caused only by difficult life circumstances. Instead, this

paper looks at whether social and environmental factors also play a role in shaping these behaviours. Behaviours such as a lack of empathy, being manipulative, and moral disengagement are not necessarily rooted only in individual psychological disorders. This paper examines whether they can also be shaped, reinforced, or encouraged by the social experiences and environments a person moves through over the course of their life. This matters for two reasons. The first is that the way we raise children, design institutions, and organize workplaces all have real consequences for how people treat one another. If these behaviours can be influenced by social conditions, then changing those conditions may help reduce or prevent them. The second is more academic, as much of the existing research on psychopathy has focused on biological and clinical explanations, and social influences have been given comparatively less attention (Hare & Neumann, 2008). This paper attempts to address that by drawing on evidence from psychology, sociology, and criminology to build a clearer picture of how the social world may contribute to behaviours commonly associated with psychopathy. The paper is organized into six sections. It begins by defining psychopathy through established clinical frameworks, particularly Hare's Psychopathy Checklist-Revised (PCL-R). It then examines three social explanations, each operating at a different level of influence. The first is social roles: how the positions people occupy can shape their behaviour toward others almost immediately, regardless of who they were before stepping into that role. The second is socialization and how the process of growing up within particular families, peer groups, and communities gradually builds the emotional and moral patterns that a person carries into adulthood. The third explanation is institutional environments and how the formal organizations people inhabit in adult life can reinforce or intensify the behavioural tendencies that roles and socialization have already shaped. The paper then briefly considers biological explanations as a counterpoint, acknowledging that social and biological factors interact rather than compete. Because this research focuses on human behaviour in social contexts that cannot be ethically or practically reproduced in a laboratory, a literature review will be used. Rather than generating new data, this paper takes findings across a wide range of existing studies and combines them to identify patterns, draw comparisons, and evaluate the strength of social explanations relative to one another. The sources used in this paper help provide a clearer understanding of how social conditions may shape, reinforce, or influence behaviours commonly associated with psychopathy.

What Is Psychopathy?

Psychopathy is not just a word for someone violent or dangerous. It refers to a specific set of personality traits that tend to appear together, such as lying frequently, manipulating others, and feeling very little guilt when hurting someone (Hare & Neumann, 2008). What makes this interesting is that people with these traits do not always stand out and in fact many people who have these traits live ordinary lives in everyday social and professional settings, and some even hold leadership positions without anyone around them noticing anything unusual (Babiak, Neumann, & Hare, 2010). The main tool researchers use to measure psychopathy is the Psychopathy Checklist-Revised, or PCL-R, developed by Canadian psychologist Robert Hare in

the 1980s. Hare made a scored checklist of 20 traits, grouped into two categories, where the first category covers interpersonal and emotional traits, such as being manipulative, lying pathologically, and lacking empathy, and the second category covers behavioural patterns, such as acting impulsively and behaving irresponsibly (Hare & Neumann, 2008). Each trait is scored from 0 to 2, which gives a maximum of 40 points. A score of 30 or above is the threshold used in research to identify psychopathy. The checklist matters here because it gives this paper a concrete and agreed-upon set of traits to examine, rather than relying on vague assumptions about what a psychopath looks like. It is also worth separating psychopathy from a similar related condition called antisocial personality disorder, or ASPD. Both involve disregarding others and breaking social rules, but they are not the same thing because ASPD focuses on persistent antisocial and rule-breaking behaviour, while psychopathy focuses more on a combination of emotional and behavioural characteristics including manipulation and a genuine absence of empathy or remorse (Hare & Neumann, 2008). Not everyone with ASPD qualifies as a psychopath, and not every psychopath has ASPD. Even if someone truly cannot feel empathy, it is tempting to assume they were simply born that way. The next three sections challenge that assumption by examining how the roles people hold, the environments they grow up in, and the organizations they inhabit may all influence whether behaviours commonly associated with psychopathy emerge or intensify.

Social Roles and Psychopathic Behaviour

A social role is the position a person holds in a group or organization; for example, a prison guard, a corporate manager, or a school president. Each of these roles comes with power over others and certain expectations about how to act. People being placed in a role with a lot of power and very little accountability can push a person toward behaviours that look a lot like psychopathy, even if they had no history of acting that way before. This is different from socialization, which shapes who a person becomes over years. Research suggests that social roles can change how someone behaves within days, simply because of the position they now hold. The most well-known example is the Stanford Prison Experiment, run by psychologist Philip Zimbardo in 1971. Zimbardo randomly assigned university students to act as either prisoners or guards in a simulated prison, and within a few days some of the students playing guards had started acting in ways that were cruel and dehumanizing toward the prisoners, and the study was forced to end after six days (Haney, Banks, & Zimbardo, 1973). However, the experiment has faced significant criticism as not all of the guards became abusive; there was variation in how different participants responded to the same role, which suggests that the role alone cannot fully explain the behaviour. Participants also specifically volunteered for a study about prison life, which means they may not represent the general population; additionally, there is evidence that the researchers gave the guards some guidance about how to maintain control, which means the behaviour was not simply because of the role alone but influenced by the researchers. Because of these limitations, the experiment gives us some reason to think that roles and situations can influence behaviour, but it is not strong enough to conclude that giving someone power will

automatically make them cruel or less empathetic. Research on real workplaces points in a similar direction. Babiak, Neumann, and Hare (2010) found that some individuals with psychopathic traits tend to do well in corporate environments, where certain characteristics such as having a willingness to compete aggressively can actually help them get ahead of others. They argued that intense competition and reward systems that prioritize individual results can create conditions where psychopathic behaviour gets rewarded rather than punished. It is worth being clear, though, that this study did not follow people before and after entering these roles, so it cannot tell us whether the environment caused these traits. What it does suggest is that certain workplaces may be more likely to attract, tolerate, or reward people who already display these behaviours. Experiments by Galinsky and colleagues found that simply feeling powerful makes people worse at reading others' emotions and less likely to consider how their choices affect others (Galinsky et al., 2006). This does not mean that powerful people are inherently broken or lack empathy; Galinsky et al. (2006) found that feeling powerful reduces people's perspective, or the ability to understand other people's viewpoints. However, the study did not directly test why this happened, and one of the possible interpretations is that people in powerful positions may feel less social pressure to consider others' perspectives, especially when there is little accountability. This suggests that the role may remove some of the social conditions that normally encourage empathic behaviour. There is also a mental side to this. Psychologist Albert Bandura identified a set of mental habits that people use to justify hurting others without feeling guilty about it, things like describing harmful actions in softer language, spreading blame across a group so no individual feels responsible, and mentally stripping others of their humanity (Bandura, 1999). These habits may help explain how people in positions of authority are able to justify certain harmful decisions, especially when the environment around them normalizes or supports those decisions. To add on, roles that reward individual success and offer little accountability can also make manipulation feel like a useful tool rather than a sign of something being wrong with the person using it.

Socialization, Peer Influence, and Social Learning

While social roles can change how someone behaves quite quickly, socialization works in a very different way; instead of being really fast, socialization is the long, slow process through which people learn how to think, feel, and relate to others by growing up in a particular environment. Socialization starts in early childhood and continues through the teenage years and into adulthood. It happens through interactions with family, friends, community, and the broader culture a person is raised in and instead of changing how someone acts in a specific situation the way a role does, socialization shapes who that person fundamentally is what they consider normal, how they respond to other people's emotions, and what moral rules they carry around with them every day and by the time someone becomes an adult, their socialization experience has already influenced everything they do, often without them even being aware of it. Albert Bandura's research on social learning is probably the most useful framework for understanding how socialization works. Bandura argued that people learn from their past experiences but more

importantly from watching and copying the people around them, especially people they look up to or spend a lot of time with such as parents and friends (Bandura, 1999) which suggests that children learn behaviours by observing what the people around them do, particularly when those behaviours are repeatedly accepted within their environment (Bandura, 1999).

Studies have also found that children who have experienced abuse, neglect, or consistent exposure to violence at home or in their community tend to develop higher levels of psychopathic traits, which are considered an early sign of psychopathic development (Frick et al., 2014). However, the study does not exactly explain how this happens, but one possible interpretation is that emotional detachment may develop as a way of coping with environments where showing vulnerability leads to harm but whether this coping response becomes a lasting personality pattern is not something the research directly tested.

Peers also play a significant role, especially during the teenage years in school. A study on violent peer influence found that teenagers with psychopathic traits were significantly more likely to behave violently when they were part of peer groups that normalized or celebrated violence, and this effect was even stronger when their own self-esteem was low (Van Zalk & Van Zalk, 2015). Similarly, if a student is in a particular friend group that says lots of slang words, the student will start to pick those words up and start to use them.

This finding is important because it suggests that adolescents with psychopathic traits can still be influenced by their social environment; in particular, the study suggests that being surrounded by peers who normalize or encourage violence may be associated with greater violent behaviour among adolescents who already display psychopathic traits (Van Zalk & Van Zalk, 2015). This challenges the idea that psychopathic traits make individuals completely unaffected by social pressure. However, the study does not show that peer groups cause psychopathic traits to develop or that they permanently change a person's personality. Instead, it provides evidence that social environments may influence how existing traits are expressed through behaviour.

Bandura also explained this through his concept of moral disengagement. He identified several ways that people learn to justify hurting others without feeling guilty, calling harmful actions nicer actions, blaming the group instead of any individual, seeing the people being harmed as less than human, and telling themselves that the victims deserved it (Bandura, 1999). These are not strategies that people come up with on their own. They are passed down through socialization, picked up from the language communities use. This process may also help explain how manipulative behaviour develops. A child who grows up watching others be manipulative to get what they want may come to see it as a normal and effective way of relating to people. A child who grows up surrounded by these ways of thinking will internalize them as completely normal. This may help explain how socialization gradually embeds harmful patterns into someone's personality, and it is important to note that not everyone who grows up in a difficult environment develops psychopathic traits; in fact, research shows that people vary in how sensitive they are to their environment, and some are more affected by harmful conditions than others (Belsky & Pluess, 2009). Biology also plays a part but this does not mean socialization is

unimportant. Even if some people are more resilient, research consistently shows that the environment still plays a significant role (Frick et al., 2014; Belsky & Pluess, 2009). For many people whose behaviour resembles psychopathy, the social world they grew up in likely played a meaningful role; it isn't the only explanation, but it is a significant one. The evidence in this paper suggests that social environments are most important in bringing out or strengthening behaviours associated with psychopathy. The case that they are the main cause of psychopathy itself is harder to make.

Institutional Environments and Behavioural Conditioning

Institutional environments are organizations that people spend lots of their lives inside, such as prisons, the military, or workplaces. It essentially works through years of repeated exposure to the norms, expectations, and pressures of a particular organization. Over time, this exposure can change how people think and behave during adulthood, through the conditioning effects of a specific institution rather than through childhood upbringing or the immediate demands of a job title.

Prisons are the most studied example of this. Research on the long-term psychological effects of incarceration has found that spending extended time in prison tends to change people in predictable ways. People who are imprisoned for long periods often become emotionally shut down, deeply suspicious of others, constantly on edge, and much less capable of the kind of open and trusting connection that healthy relationships require (Haney, 2001). These changes make sense within the prison context. Showing emotions is risky, trusting people can get you hurt, and appearing weak invites exploitation. But the problem is that these adaptations do not simply disappear when someone is released. They carry them out into the rest of their life, and these changes demonstrate that institutional environments can influence people's emotional and interpersonal behaviour over time, although they should not be interpreted as evidence that incarceration produces psychopathy. The institution has conditioned these patterns into the person over time, not through childhood and not through a single role assignment, but through years of daily exposure to a very specific social environment.

What is interesting is that the prison does not only do this to the people locked inside it. Research has also found that the officers can be similarly affected over long careers. The effects of the prison environment are not limited to prisoners, as correctional staff also work within an institution shaped by rules, power relationships, and pressures surrounding security and control. However, the evidence discussed here is stronger for the psychological effects of incarceration on prisoners than for changes in empathy among correctional officers. The prison shapes both its inmates and its staff, pushing both groups toward emotional patterns that overlap significantly with psychopathic behaviour.

The military is another clear example. Military training is purposely designed to reduce certain empathic responses, especially toward enemies, because those responses would make it extremely difficult for soldiers to carry out combat orders. Military experiences can also have lasting psychological and social effects. Litz et al. (2009) found that morally injurious

experiences such as carrying out, witnessing, or failing to stop actions that go against a person's core values can leave lasting emotional, psychological, and social damage. In other words, what soldiers are asked to do, or what they see others do, can stay with them for a long time after they leave the military. It is important to be clear, though, that this is not the same as psychopathy. Moral injury is its own distinct psychological experience (Litz et al., 2009), and the two should not be confused with each other. Research on antisocial cognition helps explain the thinking patterns behind this. A study has found that the relationship between psychopathic traits and violent behaviour is explained by the mental habits people develop, including things like seeing the world as a hostile place, minimizing the harm they cause to others, and using mental shortcuts to avoid feeling responsible for their actions (Salekin, Rogers, & Sewell, 2015). These habits are not fixed features of a disordered brain. They can be learned and reinforced by environments that consistently model and reward them. High-pressure corporate environments work in a similar way. When organizations reward harmful decisions and protect leaders from the consequences of their actions, they can gradually push people toward psychopathic patterns of thought and behaviour. (Boddy, 2011). Across all three of these institutional contexts, the evidence suggests that institutional environments can reinforce or intensify existing behavioural tendencies, though it is also likely that people enter these institutions with pre-existing traits that interact with the environment over time. Institutions can impose their own norms, reward certain behaviours, punish others, and over time reshape the people inside them. This suggests that reshaping can, in the right conditions, contribute to behaviour that looks remarkably like traits associated with psychopathy in people who had no such tendencies when they entered.

Biological Factors: A Counterpoint and Integration

It would not be fair or accurate to only present social explanations for psychopathy without also addressing the counterpoint: what biology says. Research on the autonomic nervous system found that people with higher psychopathy scores tend to show reduced physical responses when they see someone else in pain or distress (Poeppl et al., 2019). Their bodies simply react less. This points to a real biological component in the emotional detachment that characterizes psychopathy. Brain imaging research has added to this, showing that people with high psychopathy scores have reduced activity in the amygdala — the part of the brain that processes fear and empathy — and weaker connections between the areas handling emotional response and those involved in moral reasoning (Blair, 2008). Biology plays a role in psychopathy, but acknowledging that biology matters does not mean that social experience does not also matter in fact research in the field of epigenetics — the study of how your behaviors and environment can cause changes that affect how your genes work — has shown that the environment a person grows up in can actually change how their genes are expressed, alter the development of the brain, and affect how the nervous system responds to social situations, sometimes in ways that last a lifetime (Meaney, 2010). A child raised in a stressful, neglectful, or traumatic environment will often develop different neurological patterns than one raised in a safe and nurturing environment, patterns that can affect empathy and emotional processing in ways

that are hard to distinguish from those seen in clinical psychopathy. This suggests that biology and social experience can influence each other rather than acting as completely separate factors. Rather than competing, they interact throughout a person's life, each shaping how the other is expressed.

Synthesis and Discussion

Looking at all three explanations together, a picture starts to emerge. Psychopathic behaviour is not something that only comes from inside a disordered individual. Social roles, socialization, and institutional environments can all play a part in shaping, reinforcing, or intensifying behaviours associated with psychopathy. Each explanation contributes something different, social roles show that ordinary people can begin acting in ways that resemble psychopathy within days of being placed in positions of power without being held accountable, socialization shows how growing up in environments that fail to nurture empathy can slowly build patterns that persist into adulthood and institutional environments show how organizations encountered later in life can sustain and deepen those patterns over months and years. These three mechanisms work at different speeds and in different stages of life, but they all point in the same direction. Someone whose upbringing taught them to suppress emotions and distrust others will likely find it easier to take on a cold, manipulative version of a powerful role and spending years in an institution that punishes empathy and rewards detachment can reinforce patterns that socialization first put in place. Manipulation in particular appears across all three explanations: rewarded by roles, normalized through socialization, and reinforced by institutions, which makes it one of the behaviours where the social argument is strongest. Based on the evidence reviewed, psychopathic behaviour rarely appears to have one single cause; it tends to develop across a whole trajectory of social experiences. When comparing the three explanations, the Stanford Prison Experiment, which is central to the social roles argument, has faced serious criticism for methodological and interpretive problems. However, research on social roles still suggests that power and pressure can influence people's behaviour by reducing their empathy and encouraging more aggressive behaviour, although this does not show that roles create psychopathic traits. Socialization research does show associations between difficult childhood environments and psychopathic traits, but it does not fully explain how these traits develop, and some studies rely on self-reported experiences which can be inaccurate. Institutional environment research also shows that prolonged exposure to institutions can affect psychological and behavioural outcomes, but this research is observational and does not directly measure psychopathy. Overall, the evidence is strongest for social environments influencing behaviours associated with psychopathy, while evidence that they directly create psychopathy is weaker. It is also useful to think about the different ways social factors might matter. The first way is whether a social environment can actually create a behaviour or trait, contributing directly to its development. The second is whether it reinforces something that already exists, strengthening or rewarding a tendency that was already present. The third is whether it brings out a behaviour in a particular situation that a person might not show in other situations. These three possibilities do not have

equal amounts of evidence behind them; the literature reviewed in this paper shows that social environments reinforcing behaviours are associated with psychopathy, rather than directly creating psychopathy itself. Keeping this distinction in mind helps answer the extent of the research question. There are also important limitations to acknowledge about this paper. Most of the experimental work found in this review is based on roles that were completed in simulated settings not reflecting how things work in the real world. Most of the socialization research relies on self-reported findings, and most of the institutional environment research is observational rather than experimental. Additionally, it is important not to overstate the social argument, as most people who experience adverse socialization, hold powerful roles, or work in tough institutions and they do not develop psychopathic traits. Individual differences clearly matter, and biology plays a real role in moderating how people respond to social conditions. The claim being made here is not that social factors are the whole story, but that they are a necessary and an underappreciated part of it.

Conclusion

This paper set out to answer the question of to what extent behaviours commonly associated with psychopathy can be explained by social roles, socialization, and institutional environments. After examining three distinct social explanations, the answer is that social roles may contribute to ordinary people displaying behaviours associated with psychopathy relatively quickly and situationally by stripping away accountability and inflating power; socialization can gradually build the emotional and moral patterns associated with psychopathy across years of formative experience, and institutional environments can sustain and intensify those patterns in adulthood through the conditioning effects of sustained organizational exposure. Together, these explanations show that psychopathic behaviour can be influenced by social experiences, while still interacting with biological factors.

Biology is not being dismissed. The research showing that psychopathic individuals have measurable differences in their nervous systems and brain activity is real and important (Poeppl et al., 2019; Blair, 2008). Biology and social experience shape each other across a lifetime; for example, the conditions a person grows up in can alter their neurobiology, and their neurobiology can influence how strongly they are affected by social conditions. A complete understanding of psychopathy needs to account for both biology and sociology.

If social conditions contribute to psychopathic behaviour and strongly affect it, then changing those conditions can make a big difference; for example, schools and families can build environments that nurture empathy and teach children how to manage their emotions better and treat others with more care and respect, while workplaces can be designed in a way that it distributes power more fairly and hold people more accountable for their decision and in addition, prisons can be changed to reduce the psychological harm they cause to both inmates and staff.

Finally, future research should follow people over longer periods of time and pay closer attention to what happens when their social environments change. The better we understand the

social roots of psychopathic behaviour, the better placed we will be to build communities and institutions where those behaviours are less likely to take hold.

Works Cited

Babiak, Paul, Craig S. Neumann, and Robert D. Hare. "Corporate Psychopathy: Talking the Walk." Behavioral Sciences & the Law, vol. 28, no. 2, 2010, pp. 174–193.

Bandura, Albert. "Moral Disengagement in the Perpetration of Inhumanities." Personality and Social Psychology Review, vol. 3, no. 3, 1999, pp. 193–209.

Belsky, Jay, and Michael Pluess. "Beyond Diathesis Stress: Differential Susceptibility to Environmental Influences." Psychological Bulletin, vol. 135, no. 6, 2009, pp. 885–908.

Blair, R. James R. "The Amygdala and Ventromedial Prefrontal Cortex: Functional Contributions and Dysfunction in Psychopathy." Philosophical Transactions of the Royal Society B, vol. 363, no. 1503, 2008, pp. 2557–2565.

Boddy, Clive R. "The Corporate Psychopaths Theory of the Global Financial Crisis." Journal of Business Ethics, vol. 102, no. 2, 2011, pp. 255–259.

Frick, Paul J., et al. "Annual Research Review: A Developmental Psychopathology Approach to Understanding Callous-Unemotional Traits in Children and Adolescents with Serious Conduct Problems." Journal of Child Psychology and Psychiatry, vol. 55, no. 6, 2014, pp. 532–548.

Galinsky, Adam D., et al. "Power and Perspectives Not Taken." Psychological Science, vol. 17, no. 12, 2006, pp. 1068–1074.

Haney, Craig. "The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment." From Prison to Home: The Effect of Incarceration and Reentry on Children, Families, and Communities, U.S. Department of Health and Human Services, 2001.

Haney, Craig, Curtis Banks, and Philip Zimbardo. "Interpersonal Dynamics in a Simulated Prison." International Journal of Criminology and Penology, vol. 1, no. 1, 1973, pp. 69–97.

Hare, Robert D., and Craig S. Neumann. "Psychopathy as a Clinical and Empirical Construct." Annual Review of Clinical Psychology, vol. 4, 2008, pp. 217–246.

Litz, Brett T., et al. "Moral Injury and Moral Repair in War Veterans: A Preliminary Model and Intervention Strategy." Clinical Psychology Review, vol. 29, no. 8, 2009, pp. 695–706.

Meaney, Michael J. "Epigenetics and the Biological Definition of Gene × Environment Interactions." Child Development, vol. 81, no. 1, 2010, pp. 41–79.

Poeppl, Timm B., et al. "Autonomic Nervous System and Possible Determinants of Social Behavior: A Systematic Review of Its Relationship Among Psychopathy, Compassion, and Theory of Mind." Clinical Neuropsychiatry, vol. 16, no. 1, 2019, pp. 1–10.

Salekin, Randall T., Richard Rogers, and Kenneth W. Sewell. "Psychopathy and Violence: Does Antisocial Cognition Mediate the Relationship Between the PCL:TV Factor Scores and Violent Offending?" Law and Human Behavior, vol. 39, no. 2, 2015, pp. 135–146.

Van Zalk, Maarten H. W., and Nejra Van Zalk. "Violent Peer Influence: The Roles of Self-Esteem and Psychopathic Traits." Development and Psychopathology, vol. 27, no. 4, 2015, pp. 1077–1088.

Coaches, Critiques, and Confidence: A Mixed-Methods Study on Body Image in Retired Competitive Gymnasts By Bradie Hinson

Abstract

Women's artistic competitive gymnastics contains an infamous reputation for athletes' struggling with poor mental well-being, self-esteem issues, disordered eating habits, and body dysmorphia. This is assumed to come from a variety of factors in the local environment of a gymnast. Competitive gymnasts spend upwards of 30 hours a week with their coaches during practice and competitions. There is no valid research establishing coaches to be a large factor in the previously established issues. This paper responds to these challenges by measuring the impact that coaches have on gymnasts versus other additional factors. The research instruments established in the paper included qualitative interviews (n=10) and quantitative surveys (n=29). The study hypothesized that coaches hold a stronger impact than other factors such as teammates and parents, and that there needs to be regulations and education on responsible word choice from coaches to minimize impact on gymnasts. After data collection, the study concludes with example questions from both instruments that gave nearly unanimous results and describes how the pilot study was completed and analyzed. The paper affirms this hypothesis, also including respect to comparison with teammates that was shown to have a large impact. The paper concludes the findings indicate that coaches do have a large impact on the body dysmorphia and self-esteem of a gymnast.

Keywords gymnastics coaching, body dysmorphia in gymnastics, disordered eating in gymnastics, coaching relationships, teenage gymnasts

Introduction

Women's competitive gymnastics has a history of body image insecurity and negative thoughts regarding self-esteem that persists for years after gymnasts leave the sport. External influences, such as parents, teammates, and coaches contribute to these outcomes. Although there has been a shift in recent years, bringing greater awareness to the health of athletes, it remains important to emphasize how the relationship between coaches and gymnasts can influence these girls long after they have left the sport. A study completed in 2016 discusses how relationships between coaches and athletes are rarely talked about as being the catalyst for female athlete body dysmorphia and eating behaviors. They describe the relationship between coaches and athletes' body perception by conducting 28 interviews with female athletes (Beckner & Record, 2016). Understanding the insecurities and eating habits in competitive gymnastics is essential for understanding how different external factors may affect retired gymnasts, contributing to the development of eating disorders or similar mental health issues. "Each member of the gymnastics community, including gymnasts, coaches, parents, and medical professionals, has a role in shaping its culture along with the well-being, safety, and success of

the gymnasts" (Ellerbe, 2024, p.6). Understanding the role of specific influences on the gymnasts can allow for a stronger athlete with improved performance and mental well-being.

Analyzing an athlete's self-evaluation of body image can provide insight into the individuals they perceive as major influences. A study conducted by Eide (1982) used correlational research to correlate high body image to strong mental health, which is something ex-gymnasts tend to struggle with as they step into a new version of themselves. This explains the need to answer the research question, "To what extent do coaches influence body dysmorphia and perception of self for retired competitive teenage gymnasts?". The goal of this project is to determine the extent to which retired competitive teenage gymnasts attribute body dissatisfaction and lowered self-esteem to their coaches' feedback. I hypothesize that coaches have a greater influence than other factors (teammates, parents, judges, social media) due to intensive, individualized daily interactions. The study will potentially influence coaches to alter the way they phrase things to maintain positive mental and physical health for the gymnasts during and after they have left the sport of competitive gymnastics. I combined quantitative surveys (n=29) measuring perceived influence through Likert scales with qualitative interviews (n=10) capturing participants' reflections.

Coach-Athlete Relationship and Communication

An imperative part of understanding the scholarly conversation regarding coaching influence on gymnasts is fully understanding different types of relationships between coaches and athletes. In a study completed in 2013, the author deems "The coach is also a significant support source, since he/she might have the greatest influence on shaping athletes' emotions, cognition, and behaviors through day-to-day athletes" (Chen, 2013, p.1). These studies establish prior research on coach-athlete communication and interaction, however, do not correlate these findings to gymnastics or attribute influences on body image to other factors such as parents, social media, judges, teammates, etc. Coach and athlete relationships vary between each child, which is why gymnasts have thoughts and opinions regarding their own perception of self and body physique. Additionally, another study suggests that coaches should adopt a more well-rounded approach to managing their athletes, promoting a deeper understanding of what adolescent gymnasts need to maintain strong and healthy overall well-being (Kilic et al., 2024). These sources provide a deeper understanding of the varied communication and trust within coach and athlete relationships, highlighting their influence on athletes.

Insecurities and Eating Habits in Competitive Gymnastics

Gymnastics literature consistently links emphasis on leanness to disordered eating (Kerr et al., 2006; Nordin et al., 2003), but attributes causality broadly to "sport culture" rather than isolating coach feedback as the trigger. This study tests whether coaches' verbal critiques predict sustained preoccupation post-retirement. Additional examples of pressures placed on the gymnasts include plans for calorie intake and weight management for adolescent girls who need to meet the demands of a growing body. "The retired gymnasts believed an emphasis on leanness

and disparaging comments about one's body led to ongoing food and weight preoccupation" (Kerr et al., 2006, p.38). The study compares eating disorders in retired gymnasts to current gymnasts.

Likewise, a study published in the European Journal of Sport Science contradicts typical perspectives on age ranges, as gymnasts begin to feel pressure to maintain a thin ideal. They discover 92% of participants in their study demonstrate some form of body image or eating habit issues. "Athletes participating in aesthetic sports report more anorexic indices than athletes participating in either ball games or endurance sports" (Nordin et al., 2003, p.2). Overall, insecurities and eating habits that are especially prevalent among retired teenage gymnasts can stem from pressure influenced by coaches, as well as other factors. Recognizing that these challenges may persist or worsen after gymnasts leave the sport is essential to understanding the broader factors that influence body dysmorphia and self-image.

Athlete's Evaluation of Body Image

A final aspect in understanding how coaches influence athletes' body image and self-esteem is a well-rounded understanding of how athletes self-reflect on their experiences in gymnastics regarding their body image. A 2024 study hypothesized that intensity and passion in the sport are associated with disordered eating, suggesting a higher risk of disordered eating among kids in competitive gymnastics compared with recreational gymnastics. They claimed "Self-oriented perfectionism is indirectly linked to adaptive and maladaptive results" (St-Cyr et al., 2024, p.2). This demonstrates an unfortunate association between body image insecurities to strength, talent, and dedication in sports.

In a study completed by Wilczyńska et al. (2021), they hold coaches accountable for an increase of anxiety among athletes, stating "…coaches do not always use positive, empowering coaching strategies, especially under the pressure of competition." The study strives to improve how coaches react in certain situations of high pressure, with hopes of a reduction in the pressure athletes are putting on themselves to look or perform a specific way. A final study regarding how athletes put pressure on themselves to achieve an ideal body image is an article focused on investigating young men and women who value finding their ideal body. "[Women] are expected to eat small amounts of food, to care about the way they look, and to only act in appropriate and stereotypically expected ways" (Rutt, 2021). The study discovers many females are concerned about their ideal body and it affects them in their day-to-day life. It is important to understand the vast number of gymnasts affected by body dysmorphia to fully grasp who they believe their strongest influences to be as well as the extent these issues affect them.

Gap in the Research

Pre-existing research establishes a history of body image issues and self-esteem problems in the sport of competitive gymnastics. Additionally, multiple sources establish coaches as a factor influencing athletes' self-perception. However, the role coaches play in potentially spurring these issues for teenage gymnasts remains unstudied. "Disordered eating behaviors

occur in gymnastics and are often endorsed, either implicitly or explicitly, by coaches and the sport context" (Kerr et al., 2006, p.40). Although this study does hold coaches responsible for these behaviors, she also attributes other influences on these causes, including parents and judges.

The anticipated value of this project would be the education of coaches leading to the reduction of body dysmorphia and increased self-esteem, confidence, athletic performance, and body positivity. The proposed hypothetical answer to my research question is that coaches play a larger role in shaping self-esteem in teenage gymnasts than previously acknowledged; they may even be the main factor. The primary objective of this study is to evaluate the longitudinal impact of coaching communication of gymnasts' self-perception. Should the data indicate a negligible correlation, the research will pivot to identify alternative environmental or social variables that exert a more significant influence on athletes' self-concept. This study will be measured by a quantitative survey and qualitative interviews completed by retired gymnasts. The mixed-method design will allow for a measurable comparison of both statistics and quotations from these teenagers. While prior work establishes coaches as a factor (Kerr et al., 2006), no study quantifies their relative dominance or links relationship closeness to long-term body image. This project fulfills that gap through a review of scholarly sources, surveys measuring attribution rankings, and interviews.

Research Design and Methodology

Study Design

My project was completed using a mixed-method design combining qualitative interviews and quantitative survey questions. Primarily, I released a questionnaire advertised publicly through Instagram and Snapchat. The standardized questionnaire responses were based on a Likert scale, with optional responses regarding interest in being interviewed at the end. Next, I interviewed teenage girls with ages ranging from thirteen-nineteen who left competitive gymnastics to understand how they reflect on their experience with their coaches. I chose qualitative interviews based on a prior study done by Coppola et al., who selected 8 Division 1 college athletes to participate in semi-structured interviews regarding how their coaches handle sensitive topics such as body image. However, centering the interviews around gymnasts allowed for a narrower scope of data collection. Participation in the interviews was voluntary and confidential, and a consent form was sent to everyone who expressed interest in being interviewed, containing mandatory signatures from the teenagers as well as their legal guardians.

After all the data was compiled, I was able to conclude the effects of coaching word choices and how the relationships built affect body image perception and self-esteem for girls who have retired from competitive gymnastics. The combination of survey responses and interviews helped me obtain a well-rounded mix of statistics and quotations, allowing for a more complete understanding of the causes of these issues. This triangulated information allowed me to evaluate the effects of coaching on the gymnasts.

Hypothesis and Assumptions

Prior to conducting my research, I was aware that gymnasts have a reputation for dealing with body dysmorphia and declining self-esteem throughout their teenage years. Given this context, I assumed these issues may stem from interactions with different stimuli in their local environment such as parents, judges, and coaches. I wanted to explore who played the largest role in affecting the gymnast's mental well-being. By identifying the root of the issues, I assume it will be easier to educate people in the gymnastics community regarding conscious word choices and an increase in constructive criticism versus harmful speech.

In this context, I hypothesize that coaches hold greater influence on the development of body dysmorphia and lowered self-esteem in retired teenage gymnasts than previously assumed, due to the close connections and interactions developed with the gymnasts as well as their need to critique and evaluate the performance of the girls. Typically, during developmental years, competitive gymnasts foster intimate and sometimes codependent relationships with their coaches due to spending over twenty-five hours at practice with them per week. Often, the impact of it is overlooked since the fundamental relationship is simply coach-athlete. However, when considering the individualization of these relationships, it is easier to attribute coaches as a big influence even years after the gymnasts have left the sport.

Subjects

Standardized Questionnaire

The subjects who participated in my standardized questionnaire were retired gymnasts who clicked on the QR code or link through Instagram, Snapchat, and the Canvas app via a high school in Tampa. 29 girls participated in my survey ages ranging from thirteen to nineteen. This age range was selected because they recently retired from gymnastics and have had time to reflect on their experiences while still having a strong recollection. Additionally, women's gymnastics is more prevalent in the Tampa Bay area than men's gymnastics, meaning studying females allowed me to collect more data.

Qualitative Interviews

For qualitative interviews, my subjects consisted of 10 people. Interviews were the best choice for my research because they enable me to connect several aspects of the Communication Theory of Identity (CTI), such as personal, relational, enacted, and communal layers (Beckner & Record, 2016). The CTI describes the correlation between personal identity and how people communicate and interact every day (Kulper, 2023). From this, I was able to understand the meaning behind specific survey responses and understand how multiple gymnasts could be affected by the same issues but have different factors causing them. This contributed to the caliber of the mixed-method study.

I acquired these interviewees through a question at the end of my survey advertising participation in the interview. The participants remained anonymous and were labeled Subject

A-J. By interviewing each girl who participated in competitive gymnastics directly, I was able to conclude whether coaches have a strong influence on body dysmorphia and poor self-esteem.

Research Instruments

To gather data for this study, I implemented qualitative interviews and a quantitative standardized questionnaire to complete a mixed-method design allowing for a thorough answer to my research question. To begin, I contacted everyone who demonstrated interest in completing an interview through my survey. I set up interviews at local coffeehouses with them to provide a neutral space. The interview I had set up consisted of 26 questions if all were answered. By conducting these interviews, I was able to pull quotations and gain a deeper understanding of the different experiences teenage girls had in competitive gymnastics.

Table 1 depicts the questions asked during the interviews. Questions four, five, and six were only asked if the participant answered "yes" to question three. Similarly, questions eight and nine were only asked if the participant answered "yes" to question seven. Finally, consent, an overview of anonymity, and adherence to ethical practices were restated before the interview began.

INTRODUCTORY QUESTIONS
1) How long did you participate in competitive gymnastics?
2) What program were you involved in when you retired? (Optional/Compulsory/Excel)
3) This interview focuses on how coaching relationships may influence gymnasts' perception of self. Do you have one specific coach you would like to reflect on?
4) What events did this coach work with you on?
5) How long did you work with this coach?
6) What levels did you work with this coach at?
PRIMARY QUESTIONS
7) Do you remember specific things that your coaches would say to you that may have impacted how you felt about your self-esteem or body image?
8) Have you experienced challenges related to body image or similar aspects of mental well-being?
9) When do you recall these challenges beginning?
10) Did you have any significant influences or specific experiences you remember?
11) How would you describe the relationship you had with your coach?
12) Would your coach make comments discussing body physique, eating habits, or other similar topics?
13) How would you say that made you feel?
14) How did that affect your time in the gym and how you felt about going to practices?
15) Did you ever feel pressure from other people, such as parents, judges, social media, or teammates regarding body physique, performance, or something similar?
16) How would you describe your confidence in yourself while you were participating in gymnastics?
17) Do you think your coaches influenced you in a certain direction regarding your self-confidence?
18) Can you describe specific instances that a coach changed your perception of self, whether physically or mentally?
19) How would you describe internal thoughts during a typical interaction with your coach?
20) Is there anything you would like to add or share that I did not already ask?

Table 1: Interview Questions for Retired Female Teenage Gymnasts

This approach allowed for a continuation of thorough data collection allowing for analysis of the coach-athlete relationship followed by cross-references with the data compiled by the survey (Table 1).

To gather more interview participants as well as receive data from people who may want to remain anonymous, I also advertised a survey to retired gymnasts designed to gauge feelings toward coach-athlete relationships as well as body dysmorphia and self-esteem.

This survey depicted in Table 2 contained nine quantitative questions using the Likert scale, and one open-ended response at the end regarding potential interview meetings. There were two additional consent questions at the beginning and end of the survey to maintain ethical considerations. The survey was advertised towards retired teenage female gymnasts, as they are the subjects of my research. There was also an overview paragraph at the beginning of the survey regarding the purpose, anonymity, and confidentiality of the survey.

QUESTIONS	ANSWER OPTIONS
I have read the overview above, and I consent to participate in this survey	Agree
How would you classify the relationship between you and your primary coach while you were participating in competitive gymnastics?	Likert Scale
How often did comments from your coach regarding your performance influence your self-esteem negatively?	Likert Scale
How often did comments from your coach regarding your performance influence your self-esteem positively?	Likert Scale
To what extent do you believe the relationship between you and your primary coach has affected your everyday life?	Likert Scale
To what extent do you agree with this statement: My coach was the primary influence in determining how I felt about my body image and physique.	Likert Scale
To what extent do you agree with this statement: I still think about things my coaches said to me even after I retired from competitive gymnastics	Likert Scale
What was the highest level at which you competed in competitive gymnastics?	Numbers ranging from 6-10 or Equivalent
At what age did you retire from competitive gymnastics?	Numbers ranging from 13-19
Would you be open to completing an interview online or at an agreed upon location? This would take around one hour and you would receive a separate email or text regarding more information. You would also receive a free food or beverage item in compensation for your time.	Yes or no
If you selected yes, please type in your name and email. By typing this in, I understand anonymity will be compromised, however confidentiality will remain.	Open-Ended Response
I affirm that I have completed this survey to the best of my knowledge and all information provided is accurate and correct.	Agree

Table 2: Quantitative Survey Questions for Retired Teenage Female Gymnasts

This survey was designed to obtain statistical data regarding emotions and feelings towards the coach-athlete relationship in girls' competitive gymnastics, allowing for a compilation of statistical data as well as a broader outreach for survey participants.

Procedures

Flyers and Survey

Flyers containing a QR code providing the link to my survey were distributed through Instagram, Snapchat, cheerleading facilities, diving facilities, and a Canvas message at a Tampa high school. The gymnasts read the overview and answered questions regarding their interactions with coaches, as well as their thoughts about body dysmorphia and self-esteem in competitive gymnastics. The survey answers were completely anonymous, and I was able to access individual and group responses.

Once the data was gathered, it was compiled into positive and negative answers. The Likert Scale had two options with positive connotations and two with negative connotations. Then, I compared these answers with the connotation of the interviews to determine primary influences for gymnasts as well as the severity of lasting mental health effects. These groups were cross-referenced with the survey answers to determine the majority responses.

Qualitative Interview

Secondly, qualitative interviews were conducted with girls who demonstrated interest in the final few questions of the survey. There were ten participants, and the interview took about an hour to complete. The interviews were conducted to provide corroborating, firsthand evidence supporting the survey results. These interviews were also imperative to answering my research question by gathering data from a variety of coaches and different gyms, permitting a multifaceted understanding of the diverse factors contributing to body dysmorphia and poor self-esteem.

Interview data was collected by a voice recording app, which required participant approval. Additionally, notes were taken in a password-protected laptop to avoid breaches of privacy.

Ethical Considerations

To ensure my research adhered to ethical practices, I took several precautions such as ensuring and communicating full confidentiality with interview and survey responses. Participants who signed up to be interested in the interview were informed of consent and completed a form requiring their signature, as well as that of a parent or guardian. The survey included two consent questions regarding consent and truthful responses. In both the surveys and the interviews, I explicitly stated that each question was fully optional to answer, and they were aware that they could withdraw their answers at any time. My project also was approved by my school-based Institutional Review Board (IRB) to ensure I was following ethical practices.

Additionally, my survey was available to anybody who met the established criteria, ensuring random selection of participants. All responses were stored in a password-protected laptop, ensuring privacy. Overall, I made several conscious decisions to guide my research conduct using fair and ethical methods to promote accurate data collection.

Delimitations

This study includes several delimitations that specify what was and was not included, which helps maintain focus throughout the research collection. For example, one delimitation was the inability to collect data from girls currently in competitive gymnastics. As most of these girls are in a younger age range (nine-thirteen), the questions I would be asking are too sensitive for girls still interacting with their coaches every day. Additionally, these girls have not yet started reflecting on their experiences in competitive gymnastics, which might lead to responses that lack clarity.

Another delimitation in my study was the focus on interactions with coaches, excluding focus on relationships with other people in a gymnast's social circle. Excluding the other factors allowed for a deeper examination of the interconnection between the gymnast and coach. Including other factors in my research would have created a multi-variable project which would have broadened the different responses, leading to inability to cross-comparisons and inconclusive results. By focusing specifically on coaching factors, it was easier to draw results between the standardized questionnaire and qualitative interviews.

Limitations of Methodology

Prior to analyzing my data collection, it is important to discuss several limitations that affect the scope and structure of the study. The main limitation of this study was the incapability of researching girls who are currently in competitive gymnastics. Retired gymnasts were a prime group to study as they quit the sport recently and still have precise memories of interactions with coaches. Gymnasts currently participating in the sport often do not understand the serious mental and physical toll the sport can take on them. Additionally, interviewing these girls about sensitive subjects could create a negative atmosphere during practice and competitions which would be unethical and not worth the data collection

Another limitation regarding my study was that the survey remained primarily in the Tampa area. Without being able to advertise my survey directly in other cities, it was difficult to achieve a broader outreach. This presents a limitation in my findings because gyms in Tampa may have a different culture and therefore different effects on body image and mental well-being than other gyms across the country.

Results and Analysis

Quantitative Survey

The quantitative survey was released on January 7, 2026, to obtain a comprehensive understanding of gymnasts' relationships with coaches and self-esteem. The survey was advertised several ways, including social media groups, individual messaging, cheerleading facilities, and diving facilities. The questions in the survey were specialized to present a holistic understanding of the athlete, meaning every aspect of the gymnasts' experience in the sport was questioned. Overall, 29 participants responded to the survey. This was enough responses to accurately draw conclusions, considering the extremely strict qualifications needed to complete the survey. Participants in the survey must have been female, competed in the optional or excel program of competitive gymnastics (levels 6–10 or equivalent), and retired in their teenage years.

Participant Background and Gymnastics Experience

Two questions on the 12-question survey asked about participants' background in gymnastics. These questions did not require subjective answers and simply provided background that enabled an extensive analysis of the questions that followed. As seen in Figure 1, 79.3% of gymnasts retired at ages 13, 14, or 15. As stated by Keverne, a neuroscience researcher at the University of Cambridge, the prefrontal association cortex, associated with brain maturation and regulatory control of emotional behavior, begins to develop as early as age 7 and continues to develop into the age of 20 (Keverne, 2004). The age group retiring from gymnastics was at the peak of this development, leading to increased emotional involvement and reliance on individuals such as coaches, whom they saw every day. Indeed, Keverne later states, "An awareness of these sensitive phases in brain development is important in understanding how we might facilitate secure relationships and high self-esteem in our children" (Keverne, 2004). This age range indicates a high level of emotional turmoil, suggesting coaches may have seriously impacted the confidence level of these easily impressionable girls. 13.7% of respondents retired at 16 or 17, with all other respondents retiring at a younger age. It is important to understand that the age at which these gymnasts retired is essential for analyzing responses to later survey questions, including emotional responses and analyses of coaching interactions and relationships.

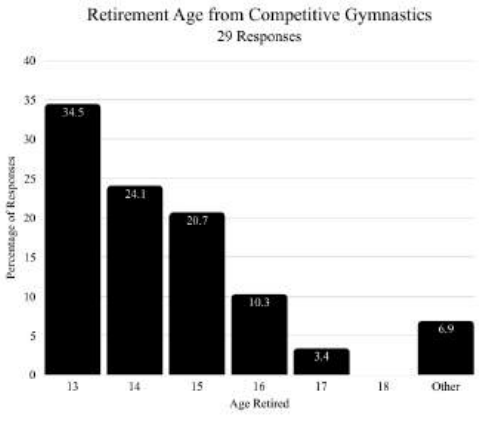
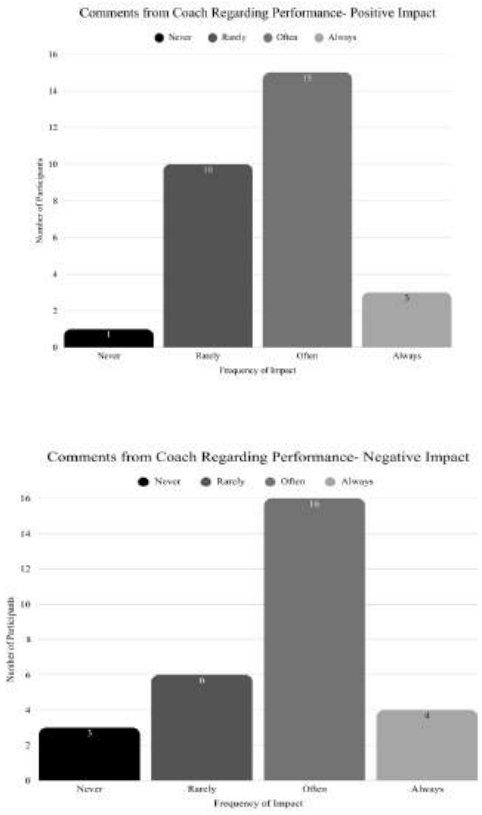


Fig 1: Average Retirement Ages of Competitive Gymnasts

Coaching Influence on Self-Perception and Well-being

Aside from confidentiality, consent, and potential interview questions, the remaining bulk of the survey focused on coach-athlete relationships and their impact on self-esteem. A strong majority of respondents (58.6%) labeled their relationship with their coach as "close," with a further majority (24.1%) considering their relationship "very close." Two questions were asked at the beginning of the survey that attempted to understand if respondents were reflecting in more of a positive or negative way regarding their experience with their coach. These questions were stated, "How often did comments from your coach regarding your performance influence your esteem positively?" The following question had the same wording, but "positively" was swapped for "negatively." These results were relatively inconclusive. As seen in Figures 2 and 3, although over half of the respondents answered "often" on both questions, there was no clear distinction between positive and negative comments regarding self-esteem. However, one conclusion clearly drawn was the overall impact the coaches had on their athletes: participants "often" or "always" had their self-esteem impacted over 60% of the time in both questions. Additionally, 72.4% of respondents still think about things their coaches said to them today, and 41.4% believe that their coach was their primary influence in determining how they felt about their body image and physique. Using this data, I was able to infer that coaching relationships can impact gymnasts for years after they have quit, exceeding typical assumptions.



Figs 2 & 3: Comparison of Coach Influence on Self-Esteem

Qualitative Interviews

At the end of the quantitative survey, participants had the option to leave a phone number or email if they were interested in being interviewed. Ten participants signed up for an interview, and these were completed over three weeks in January 2026. The survey provided one-on-one insight, including anecdotes and details that would be impossible to gauge or measure in a survey. Overall, ten interviews provided more than enough data to adequately understand individual struggles with body dysmorphia and then compare evaluations of coaching relationships. To maintain confidentiality, participants were assigned a reference letter (A–J) based on the order in which they completed the interviews.

Background and Participation in Gymnastics

To better understand the athlete's background and history in the sport, the interview included three short introductory questions. Question #3 of the primary questions included several follow-up questions if the participant chose to focus on a specific coach. Question #2 had particularly astounding results; nine out of ten participants reflected on their balance beam and floor coach. These results were not completely unexpected; balance beam and floor have a higher emphasis on clean lines, extensive dance, and artistry in the sport. In contrast, the uneven bars and vault primarily emphasize strength and power. A possible trend that was not analyzed in this project was the demographic of coaches; female gymnasts may form closer relationships with coaches who are also female or close to their age.

Support with Coaching Relationships

Five questions were asked to adequately understand how each individual gymnast reflected on their experience with their coaches. The athletes provided several anecdotes that really stood out: Participant A remembered having a scale opposing a mirror in the team locker room, and Participant A exclaimed, "There were mirrors everywhere I turned." Other participants remembered specific interactions with their coaches. Participant D shared that she often competed with girls much younger than her, and she would be called a "big." This type of language and experiences eventually led the gymnasts to get tired of this treatment. Question #5 asked how comments regarding body image and physique impacted how the girls felt about going to practice. Participant A felt "vulnerable and uncomfortable" after her coach would repeatedly tell her that "her stomach was poking out" and they "could see what she ate for lunch that day." This data demonstrates that most of the girls were not well supported in the gym and constantly felt like their appearance was being publicly judged.

Self-Perception, Body Image, and Mental Well-being

A series of questions was added to assess the range of body dysmorphia and self-esteem issues that were present after leaving competitive gymnastics. Overall, these questions provided an overwhelming consensus on consistent body image struggles. Question #2 of the primary interview questions asked, "Have you experienced challenges related to body image or similar

aspects of mental well-being?" Nine out of ten participants answered yes, prompting follow-up Question #1. Participant G was not asked the follow-up question based on her response to Question #2. Although this question had several parts, the first question ("When do you remember these challenges [see Question #2] beginning?") elicited the strongest and most significant emotional response. All participants adamantly answered right away with little deliberation. The responses ranged from different ages and grades. The most prominent was seventh grade (age 12–13), with Participants D, E, and F all remembering these starting around this time. The other answers ranged from first grade (Participant H) to ninth grade (Participant J).

Later in the interview, Question #6 asked, "Did you ever feel pressure from other people, such as parents, judges, or teammates regarding body physique, performance, or something similar?" As seen in Figure 4, about 66% of participants answered yes, with 16% attributing these issues to coaches and the remaining 50% attributing them to teammates. No participants believed these issues stemmed from judges. The consensus regarding teammate pressure appeared to stem from comparison; Participant I stated, "I mostly received pressure from my teammates because they would deflect their insecurities onto the people around them. I remember wanting to keep up with what other girls were doing which to me meant I had to look a certain way in my body." Finally, I concluded that if girls had a stronger relationship with their coaches, they were more likely to be impacted by body dysmorphia during and after their gymnastics career, which was supported by every single interviewee who considered their relationship with their coach to be close. In conclusion, these questions provided a thorough understanding of the history of body image issues and influences that began in adolescence and persisted through teenage years.

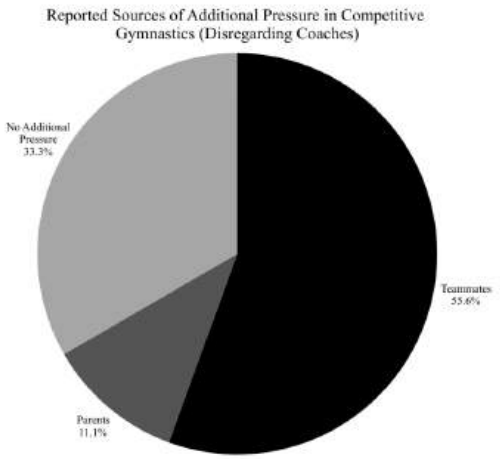


Fig 4: Body Image Pressure from Alternative Sources

Discussion

New Understanding

The goal of this project was to determine if coaches had a more detrimental impact on body perception and self-esteem than previously assumed. This was completed through data collection methods of qualitative interviews and a quantitative survey. Through this, I was able to gather sufficient data to answer my research question. This analysis reveals that, contrary to the common view that body dysmorphia and diminished self-esteem stem from multiple sources (e.g., parents, coaches, judges, teammates, and social media), these issues primarily originate from coaching influences and peer comparisons among teammates, with negligible contributions from other factors. This finding confirms the hypothesis established in the initial phase of my research process.

Fulfillment of the Gap

Previous research has already highlighted coaches as a factor in increased or decreased self-esteem. "The coaches' reactions and responses to performance can affect children's perception and evaluation of their competencies and general self-worth" (Wilczyńska et al., 2021). However, the measurable extent to which coaches impact athletes, specifically gymnasts, has left a knowledge gap in the research. I fulfilled this gap by cross-analyzing both quantitative and qualitative data over a range of athletes who competed in gymnastics at different ages, levels, and gyms to conclude that coaches do exert the greatest influence over a gymnast's perception of self after they have retired from the sport.

Additionally, my research identified a theoretical gap due to no research on the relationships each gymnast had with their coach and their subsequent body image issues. I was able to fill this gap with qualitative interviews. In every interview, if the participant believed she had an extremely close relationship with her coach, her body image issues were perceived to be worse than if a gymnast believed she was distant with her coach, where she was less impacted.

Implications

When examining the implications of this study, benefits and consequences arise for the stakeholders involved. First, coaches have a novel understanding of the impact they have on a gymnast, so they can be more mindful about their word choices regarding sensitive topics. However, the lack of evidence regarding the impact of the parents and social media can unfortunately give false ideas that they are free to make harmful and disparaging comments regarding their gymnasts' performance. This presents a call to action for a greater awareness of how the presentation of critiques can impact a gymnast, and any external stimuli regularly interacting with a gymnast should be educated about the lasting impact of their word choices. Just because coaches and teammates were shown to have most of the influence on gymnasts, the impact of other factors is not redundant.

Limitations of Conclusion

As the research was gathered and analyzed, limitations in my conclusions emerged throughout the data collection process. One limitation of my conclusions was the small sample

size, which could skew results if some gymnasts had the same coach. However, my narrow scope of research participants means it is unlikely my data would be altered if a new participant pool was used, due to the similar experiences shared by each retired gymnast. Although my participants came from different gyms, they were all located in the Tampa, Florida area, and the interviews were conducted in person. Perhaps if the study was conducted in an area that had different gymnastic cultures, the results could have been altered. Additionally, the overview explained the study's goal, so some participants may have been led to focus on their coaches despite the neutrality of the questions.

Areas for Future Research

Replication of the Current Study

To replicate the current study, another researcher must utilize the same survey and interview questions depicted in Tables 1 and 2. This will ensure exact replicability, as it uses the same questions I used to gather results. These questions were specifically designed to achieve a well-rounded understanding of the experiences of the gymnasts, and altering the questions or wording could invite bias or analysis differing from the project's intention. Also, ensure that the candidates meet the same criteria as the ones researched in my project: retired competitive female gymnasts. Most importantly, a researcher replicating my study would have to follow a similar method of analysis: comparison of quantitative and qualitative data, using anecdotes and quotations to support statistics. Finally, future researchers must adhere to the same ethical safeguards, including mandatory consent forms and overview questions, to ensure ethical compliance and data accuracy.

Different Directions

Different directions may be taken to complete a study with similar intentions. Broadening the study to include all athletes' reflections on self-esteem with coaching relationships would significantly increase the population pool, which could ensure quantity in survey and interview responses, leading to a more concrete consensus. Additionally, there are not as many sports that practice as often as gymnasts do from such a young age. This may result in more distant relationships with coaches, which could potentially skew the data more towards stimuli who see the athlete more frequently, such as increased time with parents or more time spent on social media. Another direction to take may be altering the research methods. Adding in a focus group may be beneficial to hear different perspectives add and engage with each other.

Works Cited

Beckner, Brittany N., and Rachael A. Record. "Navigating the Thin-Ideal in an Athletic World: Influence of Coach Communication on Female Athletes' Body Image and Health Choices." Health Communication, vol. 31, no. 3, 2015, pp. 364–73, https://doi.org/10.1080/10410236.2014.957998.

Chen, L. H. "Gratitude and Adolescent Athletes' Well-Being: The Multiple Mediating Roles of Perceived Social Support from Coaches and Teammates." Social Indicators Research, vol. 114, no. 2, 2013, pp. 273–85, http://www.jstor.org/stable/24720247.

Coppola, A., et al. "Coaches' Communication of Sport Body Image: Experiences of Female Athletes." Journal of Applied Sport Psychology, vol. 26, no. 1, 2014, pp. 1–16, https://doi.org/10.1080/10413200.2013.766650.

Eide, R. "The Relationship between Body Image, Self-Image and Physical Activity." Scandinavian Journal of Social Medicine. Supplementum, vol. 29, 1982, pp. 109–12, http://www.jstor.org/stable/45199668.

Ellerbe, B. "They Broke Bread with Sincere Hearts: Imagining New Gymnastics Cultures." Lateral, vol. 13, no. 1, 2024, https://www.jstor.org/stable/48826809.

Kerr, G., et al. "Disordered Eating in Women's Gymnastics: Perspectives of Athletes, Coaches, Parents, and Judges." Journal of Applied Sport Psychology, vol. 18, no. 1, 2006, pp. 28–43, https://doi.org/10.1080/10413200500471301.

Keverne, E.B. "Understanding Well-Being in the Evolutionary Context of Brain Development." Philosophical Transactions: Biological Sciences, vol. 359, no. 1449, 2004, pp. 1349–58, http://www.jstor.org/stable/4142138.

Kilic, K., et al. "A Gymnastics Coaches' Professional Development Program: Exploring a Needs-Focused Social Learning Space." Physical Education & Sport Pedagogy, 2024, pp. 1–17, https://doi.org/10.1080/17408989.2024.2335184.

Nordin, S.M., et al. "Disturbed Eating in Young, Competitive Gymnasts: Differences between Three Gymnastics Disciplines." European Journal of Sport Science, vol. 3, no. 5, 2003, pp. 1–14, https://doi.org/10.1080/17461390300073502.

Rutt, S. The Relationship between Body Image Expectations and Exercise, Eating, and Social Behavior. Trinity Student Scholarship. Trinity College Digital Repository, 2021, https://jstor.org/stable/community.34031430.

St-Cyr, J., et al. "Perfectionism, Disordered Eating, and Well-Being in Aesthetic Sports: The Mediating Role of Passion." Psychology of Sport & Exercise, vol. 73, 2024, https://doi.org/10.1016/j.psychsport.2024.102648.

Wilczyńska, D., et al. "The Effectiveness of Psychological Workshops for Coaches on Well-Being and Psychomotor Performance of Children Practicing Football and Gymnastics." Journal of Sports Science & Medicine, vol. 20, no. 4, 2021, pp. 586–93, https://doi.org/10.52082/jssm.2021.586.

Federal Architecture and the Institutionalization of Racial Risk: New Deal Housing Policy, Mortgage Finance, and the Durable Structure of Inequality (1933–1954)

By Aarav Dhar

I. Introduction

The national housing-finance system of the mid-twentieth century took shape through the federal government's response to the mortgage crisis of the early 1930s. Before the Great Depression, the mortgage system in the United States was fragmented, and there was no single federal agency to provide liquidity to the mortgage market (Harriss 7–13; Federal Home Loan Bank Act § 2). During the Great Depression, falling household incomes and property values disrupted mortgage refinancing and created hardships for both borrowers and lenders (Harriss 7–13). At the beginning of 1934, approximately half of urban houses with an outstanding mortgage were in default (Wheelock 138). Congress took action to address the crisis by passing a set of legislation in 1932, 1933, and 1934. In particular, the Federal Home Loan Bank Act of 1932 created a federal mortgage-liquidity system by establishing regional Federal Home Loan Banks connected to qualifying mortgage institutions (Federal Home Loan Bank Act § 2). The Home Owners' Loan Act of 1933 created the HOLC that was tasked with mortgage refinancing (Home Owners' Loan Act §§ 1–3). Finally, the National Housing Act of 1934 established the FHA and introduced mortgage insurance at the federal level (National Housing Act). Overall, it is evident that the measures taken by the government were not only intended to address the emergency, but also significantly changed the mortgage system by increasing the federal government's involvement. As a result, federal institutions acquired growing authority to influence how mortgages, properties, and neighborhoods were evaluated as acceptable risks.

The expansion of federal power created a historical problem that cannot be properly analyzed by reducing redlining to the influence of one agency and one set of maps. Popular accounts of redlining often position HOLC's maps as direct instructions specifying where private lenders could and could not issue mortgages. In practice, HOLC's standardized forms and maps divided urban areas into geographic security grades and recorded racial composition (Home Owners' Loan Corporation, "Blank Template" 1). At the same time, HOLC began its emergency refinancing program before completing many maps, making refinancing and security mapping chronologically and operationally distinct activities (Harriss 14–40; Hillier 210). Additionally, HOLC made emergency loans in neighborhoods it later graded D, while direct evidence that private lenders consulted completed HOLC maps in individual mortgage decisions remains limited and disputed (Hillier 210; Markley 204–08). Finally, FHA-insured lending was already concentrated away from Black neighborhoods before the relevant HOLC maps were completed, and the pattern changed little afterward (Fishback et al., "New Evidence" 1). Taken together, these observations do not make HOLC maps neutral or historically insignificant; instead, they show that discriminatory risk classification developed through a broader institutional system rather than a direct map-to-loan sequence. Considering the question in this way highlights the need to analyze a broader set of appraisal categories and methods in both federal agencies and

the private sector. Within this frame, appraisal theories circulated between private and federal actors through personnel, training, manuals, and overlapping public-private programs (Winling and Michney 56–59, 64–68). The central issue is therefore how federal agencies transformed private assumptions about risk into standardized administrative definitions and circulated those definitions through an expanding network of public and private mortgage institutions.

Answering this question requires analyzing the development of risk standards from legislative authorization through administrative practice, geographic classification, private implementation, and borrower access. The essay focuses on the period between 1933 and 1954. The former year marks the establishment of the HOLC federal emergency refinancing program, initiated by the Home Owners' Loan Act of 1933 (Home Owners' Loan Act §§ 1–4). The latter year corresponds to the statutory restructuring of the Federal National Mortgage Association's secondary-market operations, marking an important endpoint in the development of the public-private mortgage system examined here (Housing Act of 1954 § 201). Pre-1933 private practices are presented to provide context for the emergence of race-based appraisal methods. In turn, evidence from after 1954 is used to assess longer-term housing and wealth outcomes rather than extend the main institutional narrative. The essay therefore draws on primary sources, including federal statutes, FHA underwriting instructions, HOLC survey forms, executive reports, and court decisions, to understand legislative authorization, administrative practices, and contemporary assessments of barriers in the mortgage market (Federal Housing Administration, Underwriting Manual 203; National Housing Act § 203(c); Home Owners' Loan Corporation, "Blank Template" 1; President's Committee 68–69). To evaluate the relationship between administrative criteria and observed financial disparities, the essay also uses historical and quantitative studies of lending practices, homeownership trends, wealth accumulation, and segregation (Hillier 207; Fishback et al., "HOLC Maps" 1021; Aaronson et al. 1; Fishback et al., "New Evidence" 1; Faber 760–68). At the methodological level, the analysis distinguishes statutory language from administrative implementation, discriminatory appraisal standards from proof of individual lending decisions, and federal influence from sole causation. These distinctions establish the limits of the evidence and separate causes from effects. It would therefore be incorrect to claim that all D-graded neighborhoods were denied mortgage credit, that race determined every appraisal, or that federal policy alone created later spatial and wealth disparities.

Taken together, this evidence supports an institutional explanation of federal responsibility rather than either a federal-origin or private-market-only account of housing discrimination. Between 1933 and 1954, federal housing institutions did not originate racialized property valuation or dictate every individual mortgage decision; instead, they transformed preexisting private appraisal theories into standardized administrative definitions of mortgage risk, connected those judgments to federal credit support, and circulated them through a national public-private housing-finance system that distributed access to homeownership unequally and contributed to durable spatial and economic disparities. Private real-estate organizations and appraisal theorists had systematically associated race and demographic change with property

value decline well before the New Deal (National Association 5; Winling and Michney 50–53). Federal Housing Administration (FHA) underwriting manuals, Home Owners' Loan Corporation (HOLC) forms, and standardized appraisal procedures subsequently converted these private assumptions into repeatable administrative criteria (Federal Housing Administration, Underwriting Manual 203; Home Owners' Loan Corporation, "Blank Template" 1; Winling and Michney 56–59). Although HOLC security maps formed one geographic component of this broader administrative network, federal housing policy acquired broader reach through participating lenders, appraisers, homebuilders, insurance companies, veterans' loan guarantees, and secondary-market institutions (Federal Housing Administration, Fourteenth Annual Report 48; Woods, "Federal Home Loan Bank Board" 1038; Housing Act of 1954 § 201). Because FHA insurance and Servicemen's Readjustment Act loan guarantees were administered through private financial institutions operating within segregated markets, eligible nonwhite applicants could still encounter discriminatory barriers despite formal program eligibility (Woods, "Almost 'No Negro Veteran'" 393–94; President's Committee 68–69). Quantitative scholarship separately associates federal appraisal and lower HOLC grades with reduced subsequent homeownership and house values and increased residential segregation (Aaronson et al. 1; Faber 760–68). Although broader demographic shifts, local zoning, private discrimination, and regional market forces also shaped these outcomes, this multicausality does not relieve federal institutions of responsibility for granting discriminatory risk judgments, national administrative legitimacy and public financial backing. The remaining sections trace this transformation: Sections II through IV analyze the early-1930s mortgage crisis, pre-New Deal appraisal traditions, and FHA administrative codification; Sections V through VII examine HOLC geographic surveys, public-private financial networks, and FHA–VA postwar implementation; and Sections VIII through X evaluate long-term quantitative evidence, address alternative explanations, and synthesize the system's institutional legacy. The central significance of federal housing policy lies not in creating racial prejudice, but in transforming existing discriminatory assumptions into durable components of a federally supported national mortgage market.

II. Pre-New Deal Mortgage System and Housing Finance Crisis

The early 1930s mortgage crisis intensified existing mortgage distress and prompted a federal response. At the beginning of 1934, approximately half of urban houses with an outstanding mortgage were in default, demonstrating the national scale of the crisis (Wheelock 138). Farm mortgages often combined high and variable interest rates, short terms, interest-only payments, and balloon payments (Snowden 29). Although farm mortgages are distinct from urban residential mortgages, this evidence illustrates the vulnerability of one segment of the pre-New Deal credit system. Subsequently, C. Lowell Harriss concluded that private and state initiatives were ineffective, and a new mortgage program required immediate attention (Harriss 7–13). The crisis helped prompt the creation of a federal liquidity system for mortgage-lending institutions.

The Federal Home Loan Bank Act of 1932 was one of the first legislative acts to respond to the mortgage crisis by establishing a federal liquidity system (Federal Home Loan Bank Act § 2). The Federal Home Loan Bank Act organized regional Federal Home Loan Banks (FHLBs) that provided a mortgage-related liquidity system for qualifying member institutions (Federal Home Loan Bank Act § 3). The list of eligible member institutions was broad, including building and loan associations, savings and loan associations, cooperative banks, insurance companies, and savings banks within each FHLB district (Federal Home Loan Bank Act § 4). This approach is noteworthy because it is one of the first examples of federal intervention in the residential mortgage market. The system remained public-private because private lending institutions continued to originate and hold mortgages. It is essential to stress that these provisions laid the foundation for the future federal mortgage loan programs but did not specify the racial aspects of the interventions. The following federal measures, including the Home Owners' Loan Act of 1933, would define the details of the mortgage assistance and describe the target population of distressed homeowners.

The Home Owners' Loan Act of 1933 was more focused and responded directly to the homeowners' distress by creating the HOLC, which would refinance eligible mortgages (Home Owners' Loan Act §§ 1–3). Subsequently, the Home Owners' Loan Corporation (HOLC) was created as a federal corporation that provided emergency relief to distressed mortgagors by refinancing their mortgages on a long-term basis. In doing so, it is crucial to distinguish between the initial scope of HOLC's mission and its later activities (Harriss 14–40). Specifically, HOLC's 1952 report to Congress revealed that the agency had five categories of activities, including refinancing, servicing, foreclosure, management, and liquidation of properties (Federal Home Loan Bank Board 16). These categories describe HOLC's lending, servicing, foreclosure, and property-management operations; they do not identify mapping as a continuation of the refinancing process. Because the maps were created after HOLC began refinancing, their timing prevents them from explaining the corporation's earliest lending decisions. Their existence also does not prove that they determined each later mortgage decision. Nonetheless, the question of where the ideas for such maps came from still remains.

HOLC and FHA did not originate the association between racial composition and property value; that idea was already present in professional real-estate organizations such as the National Association of Real Estate Boards (NAREB). Article 34 of NAREB's 1924 Code of Ethics instructed Realtors not to introduce into a neighborhood members of a race or nationality whose presence they believed would harm property values (National Association 5). This means that the racial composition of the neighborhood was considered in property valuation. The code therefore treated racial composition as relevant to neighborhood property values. Moreover, according to the rules of the National Association of Realtors, local boards had to adopt the national rules and could be penalized for not following them (National Association 6–7). Winling and Michney describe another legacy that contributed to the HOLC's maps, namely the affiliation of Frederick Babcock with the real estate appraisal network of Richard T. Ely. Before becoming a leading FHA official, Babcock developed his views on property appraisal within

Ely's professional network, which linked demographic change to neighborhood value (Winling and Michney 50–53). Federal institutions did not invent these valuation concepts, but their forms, manuals, and programs gave the ideas wider and more standardized application. The next step is therefore to examine how related ideas entered FHA practice under the National Housing Act of 1934.

III. FHA and the Definition of Neighborhood Stability

Building upon the professional norms established in the previous section, the National Housing Act of 1934 created the Federal Housing Administration (FHA) and authorized the creation of a national mortgage insurance system. Critical to this phase is the fact that in order for a mortgage to be eligible for insurance, the project to which it pertained had to be "economically sound," a phrase that gave the Administrator wide authority in evaluating a mortgage for insurance (National Housing Act § 203(c)). In other words, while the language of the law did not explicitly invoke race, federal administrative practice had the potential to introduce racial considerations into the mortgage process. Most importantly, the way in which the Administration elaborated on this criterion is crucial, for it is this detail which reveals how the general concern for economic soundness could have led to racist housing policies. Firstly, it is necessary to examine the practices which the FHA instituted in order to establish standardized appraisal methodology.

By developing standardized procedures, the FHA created requirements which served as guidelines for standardizing appraisals around federally insured mortgages. According to the FHA's Second Annual Report, the uniformity in appraisal coming from the standardized minimum property standards was one of the key factors which determined building practices (Federal Housing Administration, Second Annual Report 15). In other words, the federal government's intervention into the mortgage market did not merely serve as a mechanism by which capital was funneled into the economy. It also unified diverse practices of local appraisal around standardized minimums and created common ways of determining mortgage risk. In essence, it fits within the larger structural change which saw varied localized phenomena become processes involving the federal government. Secondly, while the uniformity which stemmed from such procedures does not by default imply racist intent, it is worth noting that the equality of impact does not equal the equality of opportunity. It is crucial to analyze the content of those instructions which shaped appraisal practice.

As far as content is concerned, an examination of the FHA's guidance for appraisers shows that it explicitly incorporated racially charged assumptions about neighborhood stability. First, the FHA's appraisal methodology did not focus only on the home itself. Rather, it explicitly took into account future trends, particularly as they related to the neighborhood in which the home was located. The FHA's Underwriting Manual explicitly recommended restrictive covenants as a way to avoid "inharmonious use" (Federal Housing Administration, Underwriting Manual 112). Restrictions on undesirable intrusion were also explicitly advocated as a way to ensure stability. Tillotson's work contains examples of appraisal practices which illustrate that

the movement of Black people was often interpreted as an undesirable change to the neighborhood (Tillotson 35). Furthermore, other sources, including Freund's work, demonstrate that stability was often conflated with the preservation of racial homogeneity (Freund 130). This meant that discriminatory judgments did not need to originate with each individual appraiser; they could be reinforced through FHA guidance and repeated across cases. However, it is important to note that while uniformity was achieved, that does not mean that every local FHA office utilized identical criteria in all cases. Nevertheless, once such assumptions were formally placed within the guidance for appraisal, they could transcend individual appraisers and impact entire neighborhoods.

IV. Actuarial Risk Logic and Racialized Objectivity

Translating broad directives into appraisal practice required agencies to formulate specific instructions for surveyors. In implementing the directive to consider risks to the localities where mortgages were to be given, the Federal Housing Administration (FHA) created detailed instructions for its underwriters in the form of the Underwriting Manual. The guidance did not state that race alone determined mortgage risk, but it treated the presence or anticipated arrival of "inharmonious racial or nationality groups" as reducing an area's owner-occupancy appeal (Federal Housing Administration, Underwriting Manual 203). This language shows the logic through which racial composition became a variable used by federal housing officials to predict future mortgage-related risk. In other words, the association of racial composition with reduced market appeal was incorporated into appraisal practice as part of the standardized assessment procedure. This formulation does not establish that race determined every FHA appraisal. It demonstrates that racial composition was officially incorporated as a factor relevant to anticipated market appeal and mortgage risk.

The relevance of the topic can be better understood when examining the ways in which the assessments accounted for several variables simultaneously. The completed examples of the Area Description blank illustrate that surveyors followed the instructions to evaluate a wide range of characteristics from the perspective of their potential influence on mortgage security. For instance, the description of the Philadelphia C13 area from 1937 is a comprehensive list that includes socioeconomic and demographic information, features of housing stock and terrain (Home Owners' Loan Corporation, "Philadelphia C13" fields 5–6). To analyze how the interrelation between these variables affected mortgage risk assessment, one can refer to the findings of Amy Hillier, who studied HOLC grading in this city. According to her, the racial composition of a neighborhood was a statistically significant predictor of HOLC's security grade even considering physical characteristics of houses there (Hillier 207). This conclusion is particularly valuable for the discussion of the role of race as a factor influencing mortgage security from the perspective of federal authorities. While the findings do not establish race as the sole determinant, they demonstrate a statistical relationship between racial composition and HOLC grades. Because the relationship remained statistically significant after housing characteristics were held constant, physical housing quality alone did not explain the association

in Hillier's Philadelphia sample. This evidence is valuable for the analysis as it connects the discussion with the broader debates on race and urban development while illustrating the limitations of the findings. They can be used to assert the central point that the appraisal practice operationalized race as a factor influencing mortgage security without making it the sole consideration. In other words, the analysis shows that HOLC's technical apparatus incorporated racialized criteria alongside other variables rather than making race the sole consideration.

To explain why the standardized mortgage rating practices accounted for several factors when determining security, it is critical to understand the ways in which the procedures themselves gained authority. The implementation of standardized practices can be related to the efforts of the FHA to operationalize the consideration of risks in a locality, which included both race and nonracial elements. For example, Amanda Tillotson's examination of FHA practice describes how preexisting associations between race and mortgage risk were translated into standardized criteria presented as more objective (Tillotson 26). LaDale Winling and Todd M. Michney similarly argue that the FHLB Act, HOLC, and FHA incorporated preexisting valuation ideas into a new national lending regime (Winling and Michney 56–59). When taken together, these sources provide evidence for the role of standardized risk assessment in the creation of the new mortgage insurance practices. It is therefore insufficient to explain federal housing discrimination solely through HOLC grading, because HOLC did not control every mortgage decision. The FHA's independent underwriting process therefore matters alongside HOLC's neighborhood grading system.

While the standardized appraisal system considered multiple variables, the evidence examined here shows that racial composition was incorporated alongside nonracial factors in federal neighborhood-risk assessment. The findings of Fishback et al. are particularly important because they caution against treating completed HOLC maps as the source of FHA exclusion. Because FHA-insured lending was already concentrated away from Black neighborhoods before the relevant maps were completed, the two programs should not be arranged into a simple map-to-lending causal sequence (Fishback et al., "New Evidence" 1). The HOLC maps are better treated here as a distinct geographic classification system within the broader federal housing architecture. This distinction avoids assuming that mortgage originators uniformly relied on completed maps or that standardization itself eliminated opportunities for discrimination. The survey instructions and forms demonstrate how standardized tools operationalized racial information within broader assessments of mortgage security. Combined with statistical studies, this evidence supports the narrower conclusion that racial composition influenced appraisal practice without serving as the sole determinant. Other variables remained relevant, and surveyors retained room for judgment within the standardized procedure. The final point highlights why the analysis of the appraisal practices should focus on the examination of the actual maps that visualize the neighborhood risk assessment.

V. HOLC Residential Security Maps and Federal Risk Grading

Section IV discussed the ways in which federal appraisal practices created standardized procedures by introducing racial and demographic elements into risk logic. HOLC's Residential Security program extended this logic geographically by assigning security grades to urban neighborhoods. In its city-survey program, HOLC categorized entire neighborhoods rather than individual properties. The Area Description form standardized this process through fields for the city, area number, security grade, and neighborhood characteristics (Home Owners' Loan Corporation, "Blank Template" 1). The form converted varied spatial information into a standardized federal record linked to a geographic security grade. When applied to Richmond Area No. 1, the form assigned that neighborhood the lowest security grade, D (Home Owners' Loan Corporation, "Richmond Area No. 1" 1). The Area Description form therefore treated an entire neighborhood as a unit of geographic risk. With that being said, it is difficult to argue that simply giving an area a security grade 'D' resulted in HOLC denying a mortgage for every single home within this region, as the very nature of the 'composite' areas suggested that other variables could affect a mortgage application. The form's standardized approach does not by itself show that private banks followed the same classifications when deciding individual mortgage applications; it shows that HOLC treated location and neighborhood characteristics as relevant to appraisal.

In order to understand exactly how the information about a particular area was incorporated into the composite appraisal, one has to review the intermediate grades, rather than the lowest grade 'D.' For instance, when reviewing the appraisal of Philadelphia Area C13, the report highlights several positive attributes about the area, including its convenient transportation links to the city center and access to a nearby higher-grade residential neighborhood, while also recording the area's Italian population and neighborhood decline (Home Owners' Loan Corporation, "Philadelphia C13" fields 3–4). The appraisal therefore combined transportation advantages with demographic and neighborhood conditions. Read alongside HOLC's instructions, the report shows how demographic judgments entered a multifactor neighborhood appraisal without establishing that any one factor automatically determined an individual mortgage application. Amy Hillier notes that with regard to composite areas, the areas with a 'C' grade were associated with an 'infiltration,' while the areas with a grade 'D' were associated with 'undesirable' population (Hillier 217). Accordingly, Philadelphia C13 cannot establish nationwide practice, but it illustrates how HOLC combined favorable and unfavorable attributes when assigning a neighborhood security grade. By specifying exactly what HOLC accounts for when developing the composite areas and how exactly the information was processed, one is now able to move on to a more significant question. Did HOLC's composite areas affect mortgage lending, and to what extent can the map be responsible for the changes?

Assessing the historical consequences of HOLC grading requires going beyond merely describing the characteristics of neighborhoods within a given category. It involves investigating whether significant discontinuities existed across HOLC grade boundaries and whether cities subjected to federal appraisal followed different trends. To evaluate geographic discontinuities at the municipal level, Daniel Aaronson and his colleagues implemented a comparison of the

results of neighborhoods on either side of a boundary set by HOLC graders. This limits the influence of other variables, thus providing credibility to the results of their analysis. Notably, Aaronson finds that, with few exceptions, negative outcomes were not confined to the interface between D and C categories, as similar negative trends occurred at the intersection of C and B rated areas (Aaronson et al. 3). In other words, Aaronson's findings associate HOLC appraisal with disparities across several grade boundaries, not only between D- and C-graded neighborhoods. On a larger scale, comparing the progress of cities subjected to federal appraisal against the characteristics of comparable localities that were never assessed by HOLC, Jacob Faber attempted to determine whether there was a divergence in segregation trends after federal appraisal (Faber 752–60). Overall, both researchers utilize a broad perspective to analyze federal appraisal and its consequences for segregation, and their approaches complement each other. Aaronson examines discontinuities at boundaries to isolate the unique impact of HOLC, while Faber approximates the role of appraisal at the municipal level. Nevertheless, it is unclear from either study whether individual mortgage lenders actually reviewed HOLC's maps before approving a loan application. Neither Aaronson nor Faber attempted to recreate the exact decision-making algorithm used by private banks to approve or deny a mortgage application. Thus, both studies link HOLC appraisal to later negative outcomes, but neither demonstrates how particular private lenders used completed maps in individual decisions.

The geographic patterns identified by later researchers are insufficient to explain how HOLC maps were used in every completed mortgage deal. The Home Owners' Loan Corporation was initially an emergency refinancing agency, so its neighborhood appraisal activity and the production of Residential Security Maps were only a secondary activity (Harriss 14–40). HOLC's completed mortgages therefore cannot be explained simply by reference to the neighborhood maps, because the corporation began extensive refinancing before creating them. The evidence for this is provided by Amy Hillier who finds that in Philadelphia, HOLC made many emergency loans in the neighborhoods later graded D, including African Americans, Jews, and immigrants (Hillier 210). This finding contradicts the claim that HOLC categorically refused emergency refinancing in every neighborhood it later graded D, but it does not make the maps or HOLC's appraisal practices racially neutral. Nevertheless, it is insufficient to explain how HOLC's maps influenced the lending practices. HOLC's Final Report separately categorized refinancing, servicing, foreclosure, property management, and liquidation. This division reinforces the need to distinguish among the corporation's operations (Federal Home Loan Bank Board 16). Scott Markley likewise finds that discriminatory private appraisal preceded federal mapping and that direct reliance on completed maps by private lenders remains debated (Markley 204–08).

VI. Secondary Mortgage Markets and Black Financial Institutions

To understand how standardized geographic judgments circulated, it is important to identify the private institutions that could access the federal housing-finance system. The Federal Home Loan Bank Act of 1932 was developed to provide a federal source of liquidity to a range

of mortgage lenders. In this regard, Section 4 of the act allowed building and loan associations, savings institutions, cooperative banks, and insurance companies to participate if they met specific eligibility criteria (Federal Home Loan Bank Act § 4). Notably, the Federal Home Loan Bank Act did not seek to replace private mortgage banking but rather shaped the conditions for private lenders who remained at the center of the mortgage market. The Federal Home Loan Bank Act created a public-private relationship because home mortgages continued to be issued and held by private lenders. The inclusion of diverse private institutions allowed the federal government to influence the mortgage market through several types of lenders. Federal liquidity rules could therefore shape private mortgage banking without requiring federal approval of each transaction. However, statutory eligibility criteria alone do not reveal which firms actually participated, where funds flowed, or whether access was equal. The next step is therefore to examine insurance companies as institutions that originated and held federally insured mortgages.

The FHA's 1947 Annual Report reveals a significant difference between insurance companies' mortgage originations and their holdings at the end of 1947. On the one hand, insurance companies originated 15.3 percent of FHA-insured mortgages by the end of 1947. On the other hand, insurance companies held more than 40 percent of outstanding FHA-insured mortgages as of the same date (Federal Housing Administration, Fourteenth Annual Report 48). The difference suggests that originating FHA-insured mortgages and holding them as long-term assets were distinct activities. In this context, it becomes evident that private insurance companies played a critical role as a provider of long-term private capital, which in turn significantly expanded the FHA's institutional reach. Put differently, the overall value of all mortgages carried by private insurers represented a considerable amount of all FHA-insured mortgages outstanding at the end of 1947.

Large mortgage portfolios demonstrate the scale of private participation, but holdings alone reveal little about how appraisal standards and valuation practices moved between federal and private institutions. Louis Lee Woods II identifies one such avenue in the form of a joint committee. The joint committee, established in 1937, included three federal agencies and three private appraisal organizations and worked to coordinate valuation and mortgage-finance policies (Woods, "Federal Home Loan Bank Board" 1038). In turn, the very existence of this committee shows that the practices of appraisal do not exist in isolation, but rather have mechanisms which allow for their adoption across multiple organizations. Such mechanisms could influence private organizations without federal intervention in every case. Kenneth A. Snowden's historical study likewise traces the growing national organization of mortgage banking through 1940 (Snowden 1). Together, the findings of these researchers allow us to identify trends which suggest the creation of a formal channel through which appraisal practices affected multiple private organizations at the same time. However, the evidence does not establish that all private lenders participated in federal programs or that participating institutions depended entirely on federally sponsored products. While the trends described by these scholars

indicate increased uniformity, the question still remains of whether Black financial institutions were able to participate in this system on equal terms.

Participation in these public-private networks was uneven, especially for Black-controlled life-insurance companies. Michney shows that Black businesses balanced the need to provide credit to underserved borrowers, cooperate with federal agencies, and remain profitable within a segregated market (Michney 2–3). Black businesses participated in the system but did not operate on equal terms. Similar discrepancies can also be found in the analysis of HOLC's portfolio by Price Fishback and his colleagues. They show that in 1940 Black borrowers received 4.5 percent of HOLC mortgages, compared with 2.6 percent of mortgages from other lenders (Fishback et al., "New Evidence" 5). These figures show that Black borrowers were not entirely excluded and that HOLC's share exceeded that of other lenders in this comparison; they do not establish equal access. The federal housing market was therefore characterized by both participation and inequality. However, to fully understand the nature of this process, it is important to turn to the secondary market. The secondary mortgage market, in which eligible mortgages could be purchased and sold, became more formally organized through the 1954 FNMA reorganization.

The Housing Act of 1954 reorganized federal secondary-market operations and altered how institutions could participate in housing finance. Title III, §201 of the Housing Act of 1954 amended the Federal National Mortgage Association Charter Act which shaped the Federal National Mortgage Association's (FNMA) purchase and sale of mortgages "upon their origination," not the direct origination of individual home mortgages by FNMA itself. By purchasing eligible mortgages, FNMA could replenish lenders' funds and facilitate the circulation of capital through mortgage markets (Housing Act of 1954 § 201). Thus, the federal government's involvement in the insurance and mortgage programs affected the private market arrangements. The very categorization of eligible mortgages purchased by FNMA could affect the types of mortgages desired by private market lenders. This fit into the broader system described above, shaping relationships among insurance companies, appraisers, lenders, and participating Black financial institutions. The 1954 legislative changes do not of themselves demonstrate racial discrimination, differences in approval rates, or variations in the geographical distribution of mortgages acquired for FNMA. By structuring interactions among parts of the mortgage-finance system, the reorganization increased the capacity for capital to circulate and deepened ties among mortgage institutions. To understand the ways in which this happened and how it affected individual home buyers, one has to learn more about the postwar system of FHA insurance and VA-guaranteed loans, administered through private lending institutions.

VII. Postwar FHA-VA Housing and Veteran Homeownership

From the distribution of secondary-market capital to the level of individual borrowers, one can better understand how federal credit assistance advanced home-buying prospects for Americans after World War II. This was facilitated by the Servicemen's Readjustment Act of 1944, also known as the G.I. Bill, which authorized the VA home-loan guaranty program for

veterans. According to the law, a veteran could receive a guarantee of up to 50 percent of a loan against a privately issued mortgage, with the initial amount limited to $2,000 (Servicemen's Readjustment Act § 500(a)). The VA guarantee reduced private lenders' exposure to default risk and could support more favorable terms for veterans. Nonetheless, it is important to emphasize that private lenders were not required to provide credit to anybody who had served in the military. The law instead delegated rulemaking and loan-approval authority to the Administrator of Veterans Affairs (Servicemen's Readjustment Act § 504). Thus, the law itself did not describe any specific racial restrictions, but the veterans still had to rely on the private lenders' willingness to provide them with a mortgage. The statute therefore established formal eligibility without ensuring equal practical access for minority veterans. In order to analyze the impact of the G.I. Bill on these veterans, it is necessary to take a closer look at how the VA interacted with FHA insurance and how the terms of private mortgages were changed after the war.

In addition to the VA guarantee, specific financing agreements allowed federal programs to reduce barriers to obtaining homeownership. FHA insurance and VA guarantees could also be combined: Section 505 permitted a VA-guaranteed second mortgage in connection with an FHA-insured first mortgage. According to the report, the VA-guaranteed second mortgage could reach 20 percent of the purchase price, up to $4,000 (Federal Housing Administration, Fourteenth Annual Report 76). Therefore, by using both instruments, a veteran could finance a greater percentage of the purchase price through FHA-insured and VA-guaranteed mortgages. Consequently, the amount of cash needed was reduced. These arrangements did not provide a grant; both mortgages remained debts with repayment obligations. The guarantees protected private lenders against a specified portion of losses in the event of default. In the context of reduced barriers to homeownership, Daniel Fetter notes that "median down payments for VA loans were 9 percent, compared to 17 percent for FHA loans – and many VA loans had zero down payments." (Fetter 7) The ability to make smaller down payments or no down payment at all allowed many veterans to buy a home without having to save large sums of money in advance. Thus, FHA and VA financing reduced some of the financial barriers to obtaining homeownership. It is unclear to what extent the opportunities offered by the two mortgage programs were used by different racial groups. To assess whether there are disparities, it is necessary to look at the evidence of discriminatory practices on access to credit markets.

While lower down-payment requirements created opportunities, their value depended on whether eligible borrowers could actually enter the housing market. The President's Committee on Civil Rights responded to the discriminatory practices in postwar housing markets in its 1947 report To Secure These Rights. The committee recognized that minority families, like all Americans, were negatively affected by the shortage of housing. At the same time, returning veterans faced hurdles created by private real estate and community prejudices. The President's Committee on Civil Rights deemed this combination of restrictive practices as "the core of our discriminatory policy." (President's Committee 68–69) In other words, the President's Committee on Civil Rights highlighted the importance of institutional ideas that made housing discrimination systemic. FHA insurance and the VA relied on private housing industry

involvement, and the practices of this industry defined who could benefit from federal involvement. Thus, even though the governing statutes did not contain explicit racial restrictions, implementation through segregated private markets did not ensure equal access to FHA and VA benefits. The report by the President's Committee on Civil Rights therefore provides a perspective on the issue of discrimination in housing. It does not attempt to quantify the problem and define the proportion of minority households that could not take advantage of VA guaranties or FHA insurance. To understand how this systemic discrimination affected individuals, it may be necessary to review a case study of a Black veteran attempting to use these benefits.

Specifically, Collins and Margo explore the changes in the number of white and Black men who owned their homes as of 1940, 1960, and 1980. The findings demonstrate that while the percentage of white male homeowners increased by 24.2 percentage points and reached 68.1 percent, the homeownership rate among Black men rose by 18.6 percentage points and reached 44.2 percent (Collins and Margo 6–8). In other words, the homeownership gap between white and Black men increased by approximately 5.6 percentage points. Considering the trends, one cannot conclude that FHA and VA policies eliminated racial disparities in housing. However, these aggregate figures cannot determine how common experiences like Metz's were. The changes in the homeownership rates were driven by migration patterns, urbanization, and several other factors (Collins and Margo 11–13). Legal practices also contributed to the persistent gaps. In Shelley v. Kraemer (1948), the U.S. Supreme Court ruled that state courts could not enforce racially restrictive regulations, but it did not invalidate such restrictions in private contracts (Shelley 8–18). In summary, the Court barred state-court enforcement of private racial covenants but did not prohibit the private agreements themselves. Postwar FHA and VA programs expanded access to mortgage credit, while Shelley v. Kraemer (1948) altered the legal treatment of racially restrictive covenants. To assess the long-term impact of these rulings, it is crucial to analyze the overall growth of the federal housing programs and their contribution to the wealth gaps between homeowners and renters. It is also important to note the regional differences in the distribution of these programs.

VIII. Long-Term Spatial and Wealth Consequences

To assess the long-term effects of these trends, it is necessary to determine the extent to which federal programs affected postwar development. In particular, the FHA's Fourteenth Annual Report indicates how much postwar housing construction received federally insured financing by the end of 1947. The report indicates that almost one-third of new nonfarm residential construction was being financed under Title VI by the end of 1947 (Federal Housing Administration, Fourteenth Annual Report 11). Notably, this figure includes both rentals and small homes and therefore represents a large portion of postwar housing construction, not just middle-class suburban housing. It is also important to note that FHA loans do not imply direct involvement of the federal government in the construction of this housing. Instead, private lending organizations participated in the program under the protection of federal insurance. The nearly one-third figure indicates substantial federal influence on housing development beyond

the level of individual projects. By shaping the credit available to private developers, federal policy influenced postwar urban development. However, the report does not indicate what racial composition of borrowers had access to these loans, what approval criteria were used, or how evenly loans were distributed among entities. Thus, while the figure indicates the reach of the federal program, it does not provide the impact that it may have had on different groups of Americans. To understand those broader implications, it is necessary to consider federal-local redevelopment, urban renewal, and the effects of land-use planning on displacement.

Beyond mortgage insurance and construction finance, federal influence over postwar housing extended directly into urban redevelopment. Before Congress enacted the Housing Act of 1949, the President's Committee on Civil Rights addressed discriminatory housing barriers in its 1947 report, *To Secure These Rights*. The committee recommended federal action against restrictive covenants and urged stronger use of federal authority to protect equal housing access (President's Committee 169). This shows that prior to the Housing Act of 1949, federal authorities were already aware of the discriminatory housing practices and the need to intervene. However, the recommendations of *To Secure These Rights* served only as guidelines, and not as enforceable law. Two years later, Congress passed the Housing Act of 1949, which established detailed procedures for federally assisted redevelopment. Under Title I, §§ 105–106, federal redevelopment assistance depended on locally prepared plans addressing slum clearance, the disposal or reuse of cleared land, and relocation provisions for displaced residents (Housing Act of 1949 §§ 105–106). Federal policy therefore moved from investigation to direct financial involvement in local redevelopment. The expansion of federal housing policy beyond mortgage-risk administration into redevelopment was significant. However, it is still unclear to what extent it affected actual practices. To assess whether the interventions influenced the discriminatory aspects of redevelopment, such as relocations, further analysis is required.

The federal housing finance and redevelopment policies that were reviewed above in the paper had impacts that went beyond the immediate context. The focus of the analysis was on long-term neighborhood effects. In this regard, Daniel Aaronson et al. examine the role of HOLC risk classifications in neighborhood outcomes using a boundary discontinuity design and propensity-score analysis. The authors estimate that lower HOLC grades reduced subsequent homeownership rates, house values, and rents while increasing racial segregation (Aaronson et al. 1). Lower homeownership rates and house values could reduce opportunities for families to accumulate housing wealth. However, these neighborhood-level outcomes do not directly measure household net worth. While lower rent growth might be considered a positive factor, decreased values should be viewed as damaging to the ability to build wealth. An increase in segregation levels is another negative outcome of HOLC interventions. Overall, the paper illustrates that factors beyond the individual level, such as neighborhoods, can account for disparities in wealth. Notably, Aaronson et al. analyze neighborhood-level factors, so the discussion of wealth distribution is centered on the role of housing. These neighborhood-level outcomes do not directly measure generational transfers or the racial wealth gap. They should therefore be interpreted as possible mechanisms linking neighborhood conditions to wealth, not

as direct estimates of household wealth. It is important to note that the factors identified by Aaronson et al. should be considered potential contributors to the overall level of inequality rather than its sole cause. When interpreting the results of the study, it is important to compare the findings on neighborhoods with the results of studies that analyze the differences between treated and untreated cities at a larger scale.

While neighborhood boundary designs such as those proposed by Aaronson highlight disparities on either side of grading boundaries within a single municipality, assessment of longer-term spatial consequences requires consideration of entire municipalities. Addressing this broader scale, Jacob Faber compares cities that received federal appraisal with similar municipalities that were not appraised by the HOLC. Faber estimates that appraised cities experienced greater increases in residential segregation after 1930 than non-appraised cities, with multiple segregation measures remaining substantially higher in appraised cities as late as 2010 (Faber 760–68). Faber's findings at the municipal level are consistent with the argument that federal appraisal reinforced lasting spatial divisions by attaching official risk assessments to specific regions. Compared to analyses at the neighborhood level, Faber's work illustrates the connection between HOLC practices and segregation in broader contexts that extend beyond individual neighborhoods. Faber's findings at the municipal level do not account for variations in single mortgage transactions. These findings do not prove uniform local effects or establish that every lender relied on HOLC maps. Decades of segregation could influence access to mortgages, but municipal-level analysis does not address the relationship between those factors or others. An assessment of the interaction between homeownership, access to the mortgage market, and wealth accumulation would allow for a proper analysis of the consequences for individual households.

Gaps in spatial distribution are significant because the location, accessibility, and estimated value of a residence determine the returns of ownership in the long run. Examining this broader accumulation trajectory, Ellora Derenoncourt and her coauthors reconstruct the development of the Black-white wealth gap, finding that racial wealth convergence largely stalled after approximately 1950, leaving the white-to-Black wealth ratio near six to one seventy years later (Derenoncourt et al. 2). Although the stalled convergence began near the endpoint of the institutional period examined in this essay, that timing alone does not prove that FHA, VA, or HOLC policies caused it. Examining housing access over a longer horizon, William J. Collins and Robert A. Margo document that Black male homeownership rose from roughly 20 percent in 1900 to 52 percent in 1990, with most narrowing of the racial ownership gap occurring between 1960 and 1980 (Collins and Margo). This growth complicates claims of total Black exclusion, but ownership status does not ensure equal property appreciation, favorable mortgage terms, or accumulated equity. Because homeownership measures tenure rather than net asset value, expanding ownership could coexist with stalled wealth convergence; ownership rates alone do not capture differences in property appreciation, accumulated equity, or mortgage debt. Put differently, removing barriers to ownership does not guarantee equal effects on wealth, so the timing of wealth convergence and changes in homeownership should be interpreted separately.

While the policies that the federal government implemented in relation to housing are an essential factor in the consideration of the trends, neither of the papers attempts to make it the sole contributor to the wealth gap. Therefore, when evaluating the mechanisms by which the policymakers addressed the issue, it may be useful to take into consideration alternative explanations of gaps suggested by other scholars.

IX. Counterarguments

One objection to emphasizing New Deal agencies is that discriminatory appraisal practices already existed within private real estate. As established earlier, NAREB's 1924 Code of Ethics treated racial and national groups as potential threats to property values. Its national bylaws required local boards to enforce the code or face expulsion (National Association 6–7). This indicates that appraisal practices at the level of value were already organized at the national level prior to the emergence of New Deal agencies. LaDale Winling and Todd Michney show how pre-New Deal theories developed through Richard T. Ely's professional network reached the federal government via key personnel such as Frederick Babcock (Winling and Michney 50–53). The New Deal thus neither invented racial prejudice nor originated the association between race and property value. Nevertheless, private origins do not free federal programs from blame. By integrating these theories into underwriting manuals and administrative systems, federal agencies connected racialized valuation to public authority, expanding its reach, repeatability, and economic consequences. This adoption does not mean every rule derived from one source operated identically in practice. Distinguishing the private sector from the federal government requires separating HOLC's emergency refinancing operations from its appraisal and mapping activities.

HOLC security grades were based on an assessment of physical, economic, locational, credit, and demographic factors. The agency's Area Description Form standardized the collection of housing, economic, and demographic information, including racial composition and anticipated "infiltration," for use in neighborhood grading (Home Owners' Loan Corporation, "Blank Template" 1). For instance, in the case of the Philadelphia Area C13, the report notes favorable transportation access and proximity to a higher-grade area as contributing factors (Home Owners' Loan Corporation, "Philadelphia C13" fields 3–4). The same report details the occupations and approximate incomes of inhabitants, the amount of relief or aid, housing age, building types, and overall condition (Home Owners' Loan Corporation, "Philadelphia C13" fields 5–6). These data illustrate that HOLC employed a multifactorial approach to appraisal, which complicates the argument that race alone determined the HOLC security grade. Nevertheless, the inclusion of other factors does not eliminate the role of discrimination in the creation of HOLC maps. The Philadelphia C13 report records the neighborhood's Italian population and decline alongside physical and economic features, illustrating how demographic judgments entered a multifactor appraisal

The difference between the structure and the implementation also concerns the laws regulating federal housing, which relied on private organizations practicing discrimination. The

National Housing Act of 1934 required the FHA administrator to determine that the project tied to a mortgage was "economically sound" without imposing any explicit racial eligibility restriction (National Housing Act § 203(c)). Similarly, the Servicemen's Readjustment Act of 1944 allowed the Veterans Affairs administrator to approve mortgages by delegating authority to approve loans without mentioning racial limitations (Servicemen's Readjustment Act § 504). Neither statute expressly imposed racial eligibility restrictions. However, these broad delegations left substantial discretion in administration and private implementation, where discrimination could occur. For instance, the FHA Underwriting Manual warned against "socially or racially inharmonious groups" while promoting restrictive agreements as a measure against "inharmonious use" of the neighborhood (Federal Housing Administration, Underwriting Manual 112, 203). Moreover, the President's Committee on Civil Rights noted that business practices and community racism limited opportunities for nonwhite families (President's Committee 68–69). On the level of potential borrowers, Hollis Metz applied for a VA-guaranteed home loan despite having substantial income and savings but was refused by three private banks (Woods, "Almost 'No Negro Veteran'" 393–94). While the case does not represent all the veterans who served during World War II, this example demonstrates the gap between federal law and its implementation. Taken together, these sources show how formal eligibility could coexist with unequal practical access. The recognition of this inequity prompted further public calls for action.

By the late 1940s, housing policy encountered the problems of implementation from unequal application and significant legal changes, which show the lack of absolute continuity in the development of federal housing policy between 1933 and 1954. In 1947, the President's Committee reported that although FHA policy no longer actively promoted restrictive covenants, private parties could continue to impose them with impunity (President's Committee 71). The committee then called on the federal government to prohibit such restrictions and protect against housing discrimination, placing federal prohibition on the reform agenda (President's Committee 169), thus influencing the formulation of further legislative acts. Most significantly, the Shelley v. Kraemer decision of 1948 ruled that although "a contract may not by itself be unlawful or prohibited by the 14th Amendment," the enforcement of such contracts by state courts did violate the Amendment (Shelley 14–18). This illustrates the Court's attempt to intervene in the discriminatory use of private housing contracts. The Court's decision was carefully worded to indicate that while "the mere existence" of private discriminatory agreements was not prohibited by the Constitution, their enforcement by state courts did violate the 14th Amendment (Shelley 8–18). The ruling removed state-court enforcement of racial covenants but did not end private housing discrimination, demonstrating the gradual nature of reform. Shelley v. Kraemer and later legislative acts, such as the Housing Act of 1949, which formalized appropriate federal-local relations in redevelopment, and the Housing Act of 1954, which reorganized the secondary market around the FNMA Charter Act, show the changes in government housing programs without creating a complete turn towards equality (Housing Act of 1949 §§ 105–106; Housing Act of 1954 § 201). In other words, a combination of judicial and administrative measures

worked to slowly shift the legal framework surrounding federal housing policy, which did not eliminate segregated development throughout the mid-twentieth century. However, prior to concluding the analysis, one should note the limitations of the mentioned federal initiatives, which create grounds for the last argument concerning local practices, urbanization trends, and demographic shifts.

Similarly, Price Fishback and his colleagues conclude that racial bias in HOLC map creation "could account for at most 4 to 20 percent of the concentrated exposure of blacks to D neighborhoods…" (Fishback et al., "HOLC Maps" 1021). Jacob Faber likewise emphasizes that federal appraisal interacted with local institutions and regional forces, so federal policy should not be treated as the only cause of housing segregation (Faber 768–73). However, other researchers argue that the problem of unequal wealth distribution should not be entirely placed on the shoulders of one government program. Aaronson and his colleagues' study demonstrates that lower HOLC neighborhood grades "led to decreases in subsequent homeownership, house prices, and rents" (Aaronson et al. 1). These neighborhood effects are consistent with mechanisms that could contribute to later wealth inequality, but they do not establish that HOLC grading caused the nationwide halt in wealth convergence described by Derenoncourt (Derenoncourt et al. 2). Taken together, these studies support a contribution to later inequality without establishing exclusive federal causation.

X. Conclusion

The historical power of the mid-century housing system did not lie in any single statute, agency, underwriting manual, or security map, but in the connections among them. It was in the circuits of power that linked them together over time, creating feedback loops in which each element shaped and was shaped by the others. Before World War II, private real estate institutions already associated desirable population categories with desirable types of neighborhood. When the mortgage market froze during the Great Depression, federal interventions created new capacity to intervene in the selection of mortgages deemed worthy of federal backing. Through broad legislation, Congress created federal programs that operated through private institutions while pursuing public goals. Broad statutory terms such as "economic soundness" left federal agencies substantial authority to define which properties and neighborhoods qualified for support. FHA officials gave life to these directives in their underwriting manual, embedding long-standing assumptions about race, wealth, and stability into official housing policy. Through a related administrative mechanism, HOLC's neighborhood studies, Area Description forms, and security maps formalized geographic categories that combined local demographic information with judgments about default risk. Although these maps did not directly determine lending choices, they formed one part of a wider federal system of geographic risk classification. FHA regulations gained their economic leverage through a public-private mortgage market in which private mortgagees, appraisers, builders, insurers, and investors all had a role in the home-buying process. Federal risk standards could circulate

through this network and influence private choices without requiring Washington to direct each transaction.

During World War II and the postwar era, FHA insurance and VA guarantees reduced risk for private lenders, promoting homeownership among eligible buyers in ways that continued existing patterns of segregation. Even as these policies expanded homeownership opportunities, especially within white suburban markets, private institutions retained a critical role in reviewing applications and choosing locations, limiting opportunities for nonwhite buyers who encountered discrimination at multiple stages of the process. Administrative changes in the late 1940s and the 1948 Supreme Court decision Shelley v. Kraemer did not end private discrimination or FHA practices that reinforced segregation. It took several years of policy changes before legal obstacles to racial homeownership were dismantled. Nonetheless, this combination of legal, administrative, and private-market practices contributed to later segregation, with long-term consequences for the wealth of people of color. The public authorities did not need to dictate every element of the process for it to take place. They created policy through which private actors could reinforce prejudiced views about race. The authority to define the value of a home and neighborhood was shared across public and private institutions, allowing discriminatory judgments to move through several stages of the mortgage process.

The history of twentieth-century federal housing policy shows how systems of prejudice could become embedded in administrative practice through seemingly neutral terminology. Terms such as economic soundness, neighborhood stability, and mortgage security could appear racially neutral, yet administrators often defined them through assumptions about racial composition and demographic change. This approach, which federal authorities used when developing underwriting instructions, appraisal rules, standardized forms, eligibility criteria, and other guidance documents, was further reinforced by the involvement of the private sector. Once these meanings were embedded in appraisal and lending routines, discrimination could persist without explicit racial language in every transaction. At the same time, federal housing policy was financially and institutionally successful in many respects. It helped stabilize the mortgage market by standardizing loan terms and making long-term mortgages more accessible to many American families, although access remained unequal. As the history of New Deal housing policy shows, a system intended to promote economic security also turned the neighborhood itself into a unit of financial risk. Mortgage practices and federal housing policy contributed to an unequal distribution of housing resources, although they were not its sole causes. The private sector developed many of the concepts which the federal authorities later adopted at the level of national policy. However, this private origin does not remove federal responsibility for the expansion of residential segregation. The very attempt to make credit more accessible failed to account for the ways in which different groups would benefit from such policy. By integrating racialized judgments into the mortgage framework, policymakers contributed to durable patterns of residential segregation and unequal access to mortgage credit. The history of federal housing policy therefore shows how seemingly objective criteria can carry existing prejudice into durable institutions.

Works Cited

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder. "The Effects of the 1930s HOLC 'Redlining' Maps." Federal Reserve Bank of Chicago Working Paper, no. 2017-12, 2017, rev. Aug. 2020. https://www.chicagofed.org/publications/working-papers/2017/wp2017-12.

Collins, William J., and Robert A. Margo. "Race and Home Ownership, 1900 to 1990." NBER Working Paper, no. 7277, Aug. 1999. https://doi.org/10.3386/w7277.

Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. "Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860–2020." NBER Working Paper, no. 30101, June 2022. https://doi.org/10.3386/w30101.

Faber, Jacob W. "We Built This: Consequences of New Deal Era Intervention in America's Racial Geography." American Sociological Review, vol. 85, no. 5, 2020, pp. 739–75. https://doi.org/10.1177/0003122420948464.

Federal Home Loan Bank Act. Pub. L. No. 72-304, 47 Stat. 725. 1932.

Federal Home Loan Bank Board. Final Report to Congress of the United States Relating to the Home Owners' Loan Corporation. U.S. Government Printing Office, 1952.

Federal Housing Administration. Fourteenth Annual Report of the Federal Housing Administration: For the Year Ending December 31, 1947. U.S. Government Printing Office, 1948.

Federal Housing Administration. Second Annual Report of the Federal Housing Administration for the Year Ending December 31, 1935. U.S. Government Printing Office, 1936.

Federal Housing Administration. Underwriting Manual: Underwriting and Valuation Procedure Under Title II of the National Housing Act, with Revisions to February 1938. U.S. Government Printing Office, 1938.

Fetter, Daniel K. "How Do Mortgage Subsidies Affect Home Ownership? Evidence from the Mid-Century GI Bills." NBER Working Paper, no. 17166, June 2011, rev. Aug. 2012. https://doi.org/10.3386/w17166.

Fishback, Price V., Jessica LaVoice, Allison Shertzer, and Randall P. Walsh. "The HOLC Maps: How Race and Poverty Influenced Real Estate Professionals' Evaluation of Lending Risk in the 1930s." Journal of Economic History, vol. 83, no. 4, Dec. 2023, pp. 1019–56. https://doi.org/10.1017/S0022050723000475.

Fishback, Price V., Jonathan Rose, Kenneth A. Snowden, and Thomas Storrs. "New Evidence on Redlining by Federal Housing Programs in the 1930s." Federal Reserve Bank of Chicago Working Paper, no. 2022-01, 3 Jan. 2022. https://doi.org/10.21033/wp-2022-01.

Freund, David M. P. Colored Property: State Policy and White Racial Politics in Suburban America. University of Chicago Press, 2007.

Harriss, C. Lowell. History and Policies of the Home Owners' Loan Corporation. National Bureau of Economic Research, 1951.

Hillier, Amy E. “Residential Security Maps and Neighborhood Appraisals: The Home Owners’ Loan Corporation and the Case of Philadelphia.” Social Science History, vol. 29, no. 2, Summer 2005, pp. 207–33. https://doi.org/10.1017/S014555320001292X.

Home Owners’ Loan Act of 1933. Pub. L. No. 73-43, 48 Stat. 128. 1933.

Home Owners’ Loan Corporation. “Area Description (Blank Template).” NS Form-8, 1 June 1937.

Home Owners’ Loan Corporation. “Area Description C13, Philadelphia, Pennsylvania.” City Survey Files, 1935–1940, Mapping Inequality: Redlining in New Deal America, Digital Scholarship Lab, University of Richmond, 1 June 1937.

Home Owners’ Loan Corporation. Area Description (NS Form-8): Richmond, Virginia, Area No. 1. Federal Home Loan Bank Board, 3 Feb. 1937.

Housing Act of 1949. Pub. L. No. 81-171, 63 Stat. 413. 1949.

Housing Act of 1954. Pub. L. No. 83-560, 68 Stat. 590. 1954.

Markley, Scott. “Federal ‘Redlining’ Maps: A Critical Reappraisal.” Urban Studies, vol. 61, no. 2, 2024, pp. 195–213. https://doi.org/10.1177/00420980231182336.

Michney, Todd M. “Black Life Insurance Companies, Mortgages, and African American Homeownership.” Unpublished manuscript, 10 Oct. 2022.

National Association of Real Estate Boards. Code of Ethics. National Association of Real Estate Boards, 1924.

National Housing Act of 1934. Pub. L. No. 73-479, 48 Stat. 1246. 1934.

Servicemen’s Readjustment Act of 1944. Pub. L. No. 78-346, 58 Stat. 284. 1944.

Shelley v. Kraemer. 334 U.S. 1. Supreme Court of the United States. 1948.

Snowden, Kenneth A. Mortgage Banking in the United States, 1870–1940. Research Institute for Housing America, Mortgage Bankers Association, 2013.

Tillotson, Amanda. “Race, Risk and Real Estate: The Federal Housing Administration and Black Homeownership in the Post World War II Home Ownership State.” DePaul Journal for Social Justice, vol. 8, no. 1, 2014, pp. 25–52. https://via.library.depaul.edu/jsj/vol8/iss1/3/.

United States, President’s Committee on Civil Rights. To Secure These Rights: The Report of the President’s Committee on Civil Rights. U.S. Government Printing Office, 1947.

Wheelock, David C. “The Federal Response to Home Mortgage Distress: Lessons from the Great Depression.” Federal Reserve Bank of St. Louis Review, vol. 90, no. 3, pt. 1, May/June 2008, pp. 133–48.

Winling, LaDale C., and Todd M. Michney. “The Roots of Redlining: Academic, Governmental, and Professional Networks in the Making of the New Deal Lending Regime.” Journal of American History, vol. 108, no. 1, June 2021, pp. 42–69. https://doi.org/10.1093/jahist/jaab066.

Woods, Louis Lee, II. “Almost ‘No Negro Veteran . . . Could Get a Loan’: African Americans, the GI Bill, and the NAACP Campaign Against Residential Segregation, 1917–1960.”

Journal of African American History, vol. 98, no. 3, Summer 2013, pp. 392–417. https://doi.org/10.5323/jafriamerhist.98.3.0392.

Woods, Louis Lee, II. “The Federal Home Loan Bank Board, Redlining, and the National Proliferation of Racial Lending Discrimination, 1921–1950.” Journal of Urban History, vol. 38, no. 6, Nov. 2012, pp. 1036–59. https://doi.org/10.1177/0096144211435126.

The Inherent Mathematical and Philosophical Natures of Counterintuitive Decision-Making Within Society By Christopher Kim

Abstract

Individual decisions that may appear rational in isolation are able to produce outcomes that are collectively harmful when repeated across society. This paper examines this conflict through the lens of decision theory, behavioral psychology, and game theory, asking why individually rational decisions frequently produce collectively irrational outcomes and how societal incentives might better align individual and collective rationality without unnecessarily restricting autonomy. Cognitive heuristics, uncertainty, framing, overconfidence, and other biases are first considered as factors that shape how individuals perceive risk, cost, and payoff. These concepts are then applied to two cases in which private incentives can conflict with collective welfare: climate change and vaccine hesitancy. The common-resource model paints a picture of the separation between private benefit and distributed costs of greenhouse-gas emissions, while a threshold public-goods model illustrates how successful vaccination programs can create incentives to free-ride. The Prisoner's Dilemma models further identify the shared incentive structure that underlies both cases. Finally, alterations to the Prisoner's Dilemma involving rewards, penalties, repeated interactions, and the introduction of social incentives shows how cooperation can become individually rational without entirely eliminating self-interest from the model. Ultimately, this paper argues that collective irrationality does not necessarily arise from irrational individuals, but can emerge from poorly aligned incentive structures.

Literature Review

There exists a rather common assumption amongst scientists that human rationality drives social progress; it's a concept that has long underpinned both economic and political theory. Yet, mathematical and psychological research and theory persist in their challenging of this premise. One such distinct source of this tension is self-interest: when individual greed obligates oneself to prioritize momentary, short-term gain over collective stability, the resulting consequence can sometimes support productivity and exchange. However, when more resources are consumed by self-interested individuals than produced, problems begin to arise, and systemic fragility begins to show. In such cases, societal instability doesn't need to arise from malice alone, but from predictable patterns of self-interested and imperfect decision-making. This raises the central research question: Why do individually rational decisions frequently produce collectively irrational outcomes, and how can societal incentives be altered to better align individual and collective rationality without restricting individual autonomy?

For this research, decision theory and mathematics are the primary models sought to explain this phenomenon of flawed societal decision making. D. Warner North, a principal scientist/decision analyst of NorthWorks and affiliate of UCLA, supplies a simple definition: "Decision theory may be regarded as a formalization of common sense. Mathematics provides an unambiguous language in which a decision problem may be represented. In this representation,

the large and complex problems of systems analysis become conceptually equivalent to simple problems in our daily life that we solve by 'common sense."(200). In other words, it is no more than a means of logical deduction based on preferences and assumptions of the individual involved. Think of manufacturing firms. Firms often use statistical analysis, reason, and theory to weigh fixed costs of producing a component in-house against outsourcing it dependent on fluctuating customer demand and sales forecasts.

However, logic will always have its own fallacies, and for decision theory, that comes to us in the form of uncertainty. The necessity to make decisions in light of uncertainty is a core part of our lives, a part of our very being as humans. It could be stated that, indeed, uncertainty appears as the dominant consideration in many systems' problems as well as in decisions that we face in our personal lives that need to weigh advantages and disadvantages of a specific decision and its respective outcome.

This isn't to say that real human decision-making can be observed to fully follow the rules of this logical model; it can be observed that real people tend to rely on heuristics as well as cognitive shortcuts. To describe adaptive heuristics in best fashion, it is most effective to think of it as a "less-is-more" ideology. Unlike typical decision-making that requires long and strenuous calculation, adaptive heuristics is not entirely characterized with being fully rational or irrational. Rather, they are immediate forms of decision-making that, in contrast to normal decision-making, do not require computational capability nor any degrees of information (Hart 1402), like how an outfielder sprinting to catch a flying baseball simply fixes their gaze, starts running, and adjusts their speed instead of calculating wind speed or physics equations. Adaptive heuristics, thus, are often much more easily applicable when encountering convoluted strategies with noise, limited information, strategic uncertainty, etc.

In the use of adaptive heuristics, agents actively employ adjusted behavior according to circumstance based on past experience, responding to realized payoffs rather than hypothetical alternative action. But as such is the case, there must also be a measured regret, or the measured amount defined by which the payoff would have been higher had a different action been chosen then. Therefore, regret-adjusting strategies will always adjust future behavior proportional to experienced regret (Hart 1407). And because many players have the capacity to follow simple rules, the diversity in the distribution of play will still converge to the same set of correlated equilibria. In other words, non-rational individual behavior has the potential to give rise to complex rational outcomes in both the contexts of individual and societal decision-making alike.

But diversity in individual play will not always give rise to rational and productive outcomes. For example, in the case of difficult decisions interclashing with one another, and with the presence of much uncertainty, one's judgment under uncertainty is not entirely a matter of applying formal rules or calculations, but of relying on a limited number of heuristic principles. Said heuristics reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations, but they sometimes lead to severe and systematic errors (Einhorn andHogarth 3). Alongside the potential flaws of heuristics is the presence of human bias,

emotion, and instinctual misinterpretation, that of which produce holes within the decision-making model. To name a few:

1.) What humans tend to do is predict the probability of an event by the extent to which it is representative of the parent population, which in turn leads to violations of laws of probability and insensitivity to prior probabilities and sample size.
2.) A person's evaluated guess in which an event occurs is heavily influenced by the ease with which instances or associations can be brought to mind. "Availability is affected by factors other than frequency, including familiarity, salience, and media coverage" (Einhorn and Hogarth 9).
3.) People can often hold much arrogance and be overconfident in their judgments than is warranted by their collective accuracy. This neglect and inherent hubris persist even in the face of possible error.
4.) If the same problem were to have differing descriptions, it would inevitably lead to systematically different choices that are "logical" to oneself. Preferences aren't revealed by choice, and are instead manifested and created in the process of decision-making (Einhorn and Hogarth 15-18). All this said, preferences have the potential to often be intransitive and chaotically unstable.

These can all be applied to a subject we will cover later, vaccine hesitancy. A worried parent may be hesitant to vaccinate their child if witnessing another who experienced a serious reaction and treating that case as representative of what normally occurs after a vaccination, despite it revealing little about the frequency of such an outcome across the larger population. Because the reaction is vivid, emotionally tolling, and repeatedly presented through means of media, it may be recalled far more easily, distorting the parent's perception. After reading only a few articles, the parent may become incredibly certain that they have correctly determined the risk, and that it outweighs the collective benefits even though the evidence they have personally encountered is limited. And if the parent frames the reaction in two mathematically equal ways (1 out of 100 people experience outcome X as opposed to 99 out of 100 people do not experience outcome X) but emphasizes to themselves the 1 out of 100, their preference may be shaped to avoid the vaccination as to avoid the chance of outcome X altogether.

In line with chaotic preference is paternalism and cognitive bias, those of which can majorly affect the collective population's judgements and distort/influence individual decision making. Paternalism is defined as an authoritative form of control; a government, parent, or perhaps an employer that would limit or restrict a subordinate's freedom, making decisions for them that are supposedly in the subordinate's best interest, or "for their own good." What paternalism attempts at doing is restoring the conditions under which rational choice is possible. "The question is not whether to interfere, but how to do so without undermining the very agency we seek to protect" (Trout 410).

On another note, cognitive biases are unconscious errors in our thinking that can distort perception, judgment, and of course, decision-making. Being closely tied to adaptive heuristics, cognitive bias is the irrational behavior and poor outcomes that can manifest from the use of heuristics, obscuring objective reality. For ease of understanding, let us explicitly state the flaws related to the individual that directly affects the state of the community. 1. A person's guess on the frequency of an event occurring is influenced by the ease with which instances can be brought to mind. As such, occurrences like news and gossip do not produce vivid images, but keep them in public consciousness (Trout 399). 2. There are two primary biases that must be taken into heavy consideration. Firstly, the overconfidence bias: when people routinely exhibit unwarranted confidence in their judgments, even when accuracy is demonstrably low. Secondly, the hindsight bias: the removal of the otherwise natural incentive to respond to error, because what will happen will have happened anyway. It reduces the effect of the surprise, but undermines the motivation for learning (Trout 401). 3. Another one of the cognitive biases we must understand is framing. Framing is the process, whenever a problem is presented, that influences the selections of possible decisions. By shaping what appears reasonable, the individual can create connotations that are positively or negatively associated with the decision or impending outcome, rather than immersing themselves in the supposed rationality behind it all. It's of importance to note that paternalism and cognitive biases aren't simply theoretical curiosities. They have costs to each individual chooser, to society, and to public institutions. These biases help explain to us why individually "rational" decisions can be based on distorted perceptions of risk, probability, and payoff. When these distorted judgements are repeated across a population, they contribute to the collectively irrational outcomes that will be examined later in specific cases.

Methodology

All aforementioned ideas can be collectively conceptualized and modeled through various game theory mechanics, as well as numerous dilemmas that create and manipulate game models (Swedberg 316-18) of decision theory to best interpret the circumstance the human race finds itself in, which is the massively influential and persistent existence of self-interest and its effects on the collective population's choices. In order to best model this circumstance, as well as its solutions, we must establish some assumptions. The primary factor of said game models should be considered as such: people are entirely capable of deciding and have the right to decide what they want. Additionally, there should be no question of "forcing a good on a person, or forcing on him more than he puts in a claim for". It's also important to note that "Liberty and order could be achieved by setting up rules that 'guide,' 'enlighten,' but do not 'stamp out' self-interest, which is the 'motor of society" (Gališanka 302-03). Although normally associated with a negative connotation, self-interest is a fundamental driving force that prioritizes one's survival, motivation, and potential cooperation above all else. Without it, there wouldn't be individuals that could actively seek to fulfill their basic needs, including their fulfilling of their desires to specialize in certain skills or engage in exchanges that create a functioning society.

This brings us to the investigation of the research question. The following analysis will first examine climate change and vaccine hesitancy as two cases in which individually reasonable decisions can produce collectively harmful outcomes. Incentive structures will then be modeled and compared through common-resource, public-goods, and Prisoner's Dilemma frameworks to identify mathematical similarities that are underlying otherwise different social problems. To conclude, the analysis will consider whether these undesirable equilibria can be altered through changes in rewards, costs, repeated interactions, and social incentives, while returning to the question of how far society may legitimately influence individual choice.

Let us first take a look at the case of climate change. Nearly every actor within our society faces the incentive that encourages a consistent reliance on inexpensive fossil fuels (such actors include individual consumers, sovereign nations, varied corporations, etc.). Businesses are capable of maximizing profit by minimizing production costs. Consumers are always in pursuit of affordable and convenient transportation and electricity. Governments typically prioritize short-term economic growth over long-term environmental sustainability because electoral cycles reward immediate prosperity, much rather than future stability. Within the present, this would be considered "rational." However, as a result, science indicates greenhouse gas emissions continue to accumulate beyond sustainability, producing rising temperatures, extreme weather, and ecological instability as a consequence (Intergovernmental Panel on Climate Change). This in turn harms everybody, including the aforementioned primary greenhouse gas contributors. Regardless of this fact, if we choose to analyze each of these small facets of society independently, none of these decisions seem to be irrational. Each individual entity simply responds logically to the incentive and immediately attempts to maximize utility according to their own interests. This correlates to what our decision theory predicts: rational agents choose from given alternatives that maximize their expected payoff, given that there is sufficient available information and personal preference (in which case there is). As such, the problem shouldn't arise from irrationality at the individual level. Instead, it arises from the interaction between millions of individually rational decisions occurring simultaneously. It is fundamentally a clashing of wills and desires.

Earth's atmosphere functions as a shared resource in which every participant involved benefits from emitting carbon whilst the environmental costs are dished out and distributed fairly evenly across the global population. Additionally, since each individual polluter would contribute only a small fraction of global emissions, reducing one's own would appear to have a negligible impact in the pursuit to dampen impact on the environment. The only consequence of this is the imposing of significant economic costs. Every participant possesses an incentive to continue polluting while also hoping that others will voluntarily reduce their emissions, despite the amount of net profitability stemming from the usage of fossil fuels being a very easy, very obvious gold mine. Taking into consideration aspects of game theory, this can create an equilibrium in which continued pollution becomes a stable strategy despite the fact that everyone recognizes universal cooperation would create far greater collective benefit. It is Pareto inefficient: all participants *could* be better off through coordinated emission reduction. Yet, there

is no sufficient number of individuals with sufficient enough incentive to initiate cooperation properly.

A simplified common-resource model can mathematically represent this. Let e_i represent the emissions produced by one player and let E represent total emissions across all players. If $B(e_i)$ represents private economic benefit whilst $D(E)$ represents collective environmental damage, then individual players' payoffs can be represented as the following equation:

Model 1
Simplified Common-Resource Model of Climate Change

$$U_i = B(e_i) - D(E)/N$$

Note. Author-created illustrative model with assistance from OpenAI's ChatGPT; values are hypothetical and are not empirical estimates.

where N is representative of the number of players sharing the environmental cost. But in contrast to the individual, society itself must experience the full environmental damage $D(E)$. Hence, a specific action can remain individually profitable even when its total social effect is negative.

Humans also tend to systematically value immediate rewards much more heavily than future consequences. Due to the fact that 1.) industrial production and 2.) inexpensive energy is realized immediately, people naturally undervalue future environmental harm. There also exists the availability heuristic that causes people to underestimate the exact severity of gradual climate change because of the presence of other immediate, overwhelming economic concerns such as employment or inflation. Consequently, no individual would like to take any form of responsibility or action for the harm done against the collective.

An additional exemplary case regards vaccine hesitancy within a large group. Modern vaccines function in a way that protects both the individuals who receive them and all else who are unable to receive them by producing herd immunity, the state in which a sufficient number of members of a population are immune to prevent widespread transmission. If vaccination rates can remain sufficiently high, contagious disease struggles to spread, hence indirectly protecting those of the population who could not be vaccinated due to varied medical conditions or complications. Putting this factor into perspective, it could seem logical to claim that widespread vaccination is the optimal decision that would produce the optimal outcome. However, game theory already demonstrates to us that the incentives facing a number of individuals are not necessarily aligned with the collective ideal.

If vaccination rates increase, the probability of encountering a disease would decrease significantly accordingly. As an unforeseen consequence, this perceived success changes the incentive structure itself. If nearly every member of a society receives the vaccination, an individual who thinks of themselves as clever could reason that they can safely avoid the most

minimal of inconveniences, costs, or probable risks associated with receiving a vaccination and yet still benefit from the protection generated by everyone else's participation. This free-rider incentive has also been observed experimentally, particularly whenever individuals perceive that others' vaccination already provides them with indirect protection (Betsch et al. 978-985).

The phenomenon of vaccination hesitancy closely resembles and hence can be modeled as a threshold public-goods game in which costly individual vaccination contributes toward a collective level of protection (Lim and Zhang). Every participant would gain if everyone contributes, but each individual would gain even more if others contribute while they withheld their own contribution. As such, the resulting Nash equilibrium would often contain fewer contributors than would be socially optimal. Suppose a hypothetical population consists of ten individuals and that population-level protection must require at least eight of those ten to cooperate in receiving vaccinations. If eight other individuals are vaccinated, an additional individual can abstain while continuing to receive the collective benefit, making free-riding the individually attractive decision. However, say only seven others are vaccinated. The same individual's participation then becomes decisive. Cooperation pushes the population above the required theoretical threshold, whereas defection allows the public good to fail. The payoff associated with vaccination changes itself according to the already existing level of cooperation. The greater the success of widespread vaccination, the less necessary one individual's contribution will be, allowing the system's success to generate an incentive to free-ride.

Table 2
Simplified Threshold Public Goods Model of Vaccine Hesitancy

# of other individuals vaccinated	Player Vaccinates	Player Does Not Vaccinate	Individually Rational Choice
8 - threshold reached	$4 - 1 = 3$	4	Free-ride
7 - player determines whether threshold is reached	$4 - 1 = 3$	-5	Vaccinate

Note. Values are hypothetical payoff rankings rather than measurements. A collective benefit of 4 is produced once at least eight individuals are vaccinated, while vaccination carries an individual cost of 1. The value of -5 represents the hypothetical loss associated with failure to reach the population-protection threshold, reflecting increased risk and cost of disease transmission. When the threshold has already been reached by other players, the individual receives a greater payoff when free-riding. However, when the individual's participation determines whether the threshold is reached, vaccination becomes the rational strategy. Author-created illustrative model with assistance from OpenAI's ChatGPT; values are hypothetical and are not empirical estimates.

Herd immunity describes a state in which immunity within a population reduces the risk of infection even among susceptible individuals who are not themselves immune (Fine et al. 911). As more individuals choose to not receive vaccinations, herd immunity weakens until disease outbreaks inevitably become likely. The conditions that originally made non-vaccination seem rational completely dissipates as more people adopt similar reasoning that neglects vaccination. And because individuals rarely ever calculate these probabilities, the adaptive

heuristics described earlier become important in that people simplify their decision by relying upon immediately accessible experiences, stories, rumors, and risks.

A plethora of other psychological biases could also strengthen this deadlocked Nash equilibrium. One particular influential bias is the availability heuristic. Despite their rarity, serious adverse reactions to vaccines tend to receive disproportionate media attention due to their emotional vividness and how widely circulated it is through social media and news outlets (Einhorn and Hogarth 9; World Health Organization, "Vaccines and Immunization"). Additionally, diseases such as polio or measles have become relatively uncommon because of the success of vaccination programs (Centers for Disease Control and Prevention; World Health Organization, "Poliomyelitis"). As a consequence, many can readily imagine stories regarding vaccine complications whilst having very little firsthand experience or knowledge on the diseases vaccines actually prevent. Hence their perception of risk becomes distorted. It leads them to overestimate the dangers of vaccination and underestimating the greater risk associated with remaining unvaccinated.

A kind of psychological bias, or cognitive bias, is the omission bias, where individuals perceive harm resulting from inaction as drastically preferable to harm resulting from action, despite the fact statistical probabilities are identical. For example, a parent may perceive an extremely rare vaccine complication as more distressing than their child contracting a preventable disease because the former follows a deliberate decision whereas the latter follows passive inaction. On an adjacent note, confirmation bias can encourage individuals to seek information supporting their skepticisms while dismissing contradictory evidence, which is especially easy to find within online communities or social media. Framing would also alter this calculation: presenting vaccination just in terms of its possible adverse effects can and will produce a different response than presenting the same decision in terms of the diseases and complications it can prevent.

These psychological mechanisms all describe how an individual perceives a decision. In the midst of this, game theory becomes a necessity when those individually perceived choices begin interacting with the choices of everyone else. Where decision theory asks which option would appear most rational to a single chooser, game theory asks how that choice can change when its payoff depends upon the strategy selected by another. If we decide to turn to textbook game theory, a simplified Prisoner's Dilemma matrix can demonstrate how contradictions can arise mathematically. Suppose that two individuals, Player A and Player B, each possess the choice to either cooperate or defect. If both cooperate, each receive a payoff of 3. If one defects while the other cooperates, the defector receives a payoff of 4 whilst the cooperator receives only 1. If both defect, each receives a payoff of 2 (numbers do not represent measurements or quantities; rather, they rank desirability of each outcome i.e. $4 > 3 > 2 > 1$).

Table 3
Simplified Prisoner's Dilemma Payoff Matrix

	Player B cooperates	Player B defects

Player A cooperates	(3, 3)	(1, 4)
Player A defects	(4, 1)	(2, 2)

Note. Payoff values are hypothetical rankings of outcome desirability rather than empirical measurements. Mutual cooperation produces a superior collective outcome, while the individual incentive to defect leads both players toward an inferior Nash equilibrium. Author-created illustrative model based on the standard Prisoner's Dilemma framework.

Within the matrix, the first number of each pair represents Player A's payoff and the second number represents Player B's. For either player, it would appear that defecting is simply the superior strategy. If the other player cooperates, the defecting player receives a payoff of 4 rather than 3. If the other player defects, the defecting player still receives a payoff of 2 rather than 1. Hence, both players should be naturally drawn towards mutual defection and receive payoffs of 2 and 2, regardless of the fact that mutual cooperation clearly produces a superior outcome for them both. Henceforth the game is in a state of Nash equilibrium: neither player can improve their position by changing their strategy. However, the equilibrium is objectively inferior for the collective, telling us that equilibrium and social optimality are clearly not the same thing. And this is reflected in the broader equilibrium principle that was discussed earlier. Individual strategies can converge toward a stable collective pattern without said pattern being rational or desirable for a society as a whole.

Looking back at vaccination and climate change, this relatively simple model is applicable. In the case of vaccination, "cooperation" between parties represents becoming vaccinated and contributing towards protection of the total population, while "defection" would represent remaining unvaccinated and seeking to benefit from the protection created by others without contributing anything to the system yourself. In the climate-change case, "cooperation" represents choosing to reduce greenhouse-gas emissions despite potential money loss associated with doing so, whereas "defection" represents continuing to emit greenhouse gases while depending upon others to reduce their own emissions, therefore stabilizing the system. Although there exist many other extraneous factors not taken into account, the simplified model reveals that there is one primary incentive problem shared by both of the aforementioned cases: each player can gain individually if they allow others to bear the cost of cooperation.

Discussion

But what about potential solutions? If individually rational decisions could converge towards collectively undesirable consequences, then simply asking people to behave less selfishly would provide an incomplete solution. The problem isn't necessarily rooted in self-interest itself, but rather from an incentive structure in which selfish behavior favors defection over cooperation. The structure of the game rewards players for behaving in a manner that is contrary to the collective interest. A more durable solution has to therefore alter the game and its boundaries itself, making the individually rational choice increasingly compatible with

the collectively rational one. Yet, deliberately altering individuals' incentives immediately returns to us the earlier problem of paternalism. Just how far may institutions guide individual choices in the pursuit of a collectively preferable outcome?

Suppose society introduces an additional reward (let's call it variable r) for cooperating and some sort of penalty or additional cost (p) for defecting. The cooperation payoff now becomes $(3 + r)$ when the other player cooperates and $(1 + r)$ when the other player defects, whilst the payoff from defection becomes $(4 - p)$ and $(2 - p)$. In order for cooperation as a strategy to become more attractive to either player, the combined effect of the reward and penalty must satisfy one simple condition: $r + p > 1$. Society doesn't need to eliminate self-interest or convince all individuals to become altruistic entirely. If society can restructure the rewards and costs surrounding behavior, cooperation can become the strategy that better serves the individual.

One such way this can be done is solely by increasing the personal reward associated with cooperation. In the case of climate change, environmentally beneficial choices must have players accept some sort of immediate economic or practical cost while said environmental benefits are distributed throughout society. Governments and various institutions can alter this imbalance by making it so that lower-emission alternatives are less costly, prompting incentive for cleaner technologies, or otherwise increasing the economic attractiveness of reducing emissions. A corporation that previously gained more by continuing high-emission practices can now gain more by transitioning to cleaner ones once financial and technological incentive changes.

Table 4
Simplified Prisoner's Dilemma with Reward for Cooperation

	Player B cooperates	Player B defects
Player A cooperates	$(5, 5)$	$(3, 4)$
Player A defects	$(4, 3)$	$(2, 2)$

Note. This altered payoff matrix adds on a hypothetical reward r where $r = 2$ to cooperative choices whilst leaving the payoffs for defection unchanged. Cooperation then becomes the individually preferable strategy when the reward is sufficiently large to overcome the original advantage of defection (specifically, $r > 1$). Author-created illustrative model based on the standard Prisoner's Dilemma framework.

This principle can be applied through the opposite mechanism as well. Instead of increasing the reward for cooperation, society can increase costs associated with defection. This is especially relevant when a player receives a private benefit while distributing much of the resulting harm to others (in the case of climate change, through environmental externalities). An individual, institution, or nation can receive benefits of inexpensive energy or industrial activity while merely absorbing a small portion of the environmental cost created by its own emissions. Additionally, continued emissions may remain rational if the player's calculations include only

personal benefit and personal cost. But by forcing the player to bear a greater proportion of the costs generated by their behavior, institutions can change this calculation for a greater good.

Table 5
Simplified Prisoner's Dilemma with Penalty for Defection

	Player B cooperates	Player B defects
Player A cooperates	$(3, 3)$	$(1, 2)$
Player A defects	$(2, 1)$	$(0, 0)$

Note. This altered payoff matrix imposes a hypothetical penalty p where $p = 2$ on defecting choices while leaving the payoffs for cooperation unchanged. Like Table 4, cooperation becomes the individually preferable strategy when the penalty is sufficiently large to overcome the original advantage of defection ($p > 1$). Author-created illustrative model based on the standard Prisoner's Dilemma framework.

Vaccination presents to us a different application. Cooperation can also be encouraged by reducing the costs associated with participating. But while in both of these games players simply weigh the costs of cooperating against defecting, players considering receiving a vaccination do not gain a comparable productive benefit from "defection." Instead, players are avoiding what they perceive to be the cost of cooperation, like time, inconvenience, possibility of side effects, etc. Making vaccination easier to obtain, reducing logistical barriers, and improving access to reliable information can all modify the perceived payoff associated with cooperation. An individual that once may have perceived vaccination as requiring substantial effort might evaluate the same decision slightly differently after costs are reduced, hence altering which strategy appears rational.

Society can also engineer a different solution by putting the game into the perspective of a repeated game. The original matrix / Prisoner's Dilemma assumes that the players involved choose once and then never interact with the game again. But real societies rarely ever operate under such conditions. Nations can negotiate repeatedly, institutions engage with markets, individuals remain members of communities, governments depend on each other, etc. The moment future interactions become relevant, individuals must not only consider the immediate payoff of defection, but also the off-chance that present behavior will affect future cooperation, therefore practicing regret-adjusting behavior in order to modify future strategies according to whether cooperation or defection will favor their needs most. For instance, an individual may gain an additional unit of payoff by defecting today, but this short-term advantage becomes less attractive if defection prompts other players to stop cooperating tomorrow, thereby incentivizing said individual to stop defecting assuming it is the decision with the less favorable outcome in the longer-term. This is why factors like reputation and knowledge of initial behavior of other players shape the game in ways that make them differ from single instance games.

Repeated interactions are highly important for climate change because countries must make decisions while being uncertain about other nations' intent to fulfill their own

self-interested commitments. A government can be willing to accept economic costs associated with reduced emissions if it believes that its competitors will also make similar sacrifices, yet hesitates if it expects others to continue emitting freely. In such circumstances, the failure to cooperate may be caused not just by selfishness, but also by distrust. So, if countries can observe whether others are fulfilling their commitments and expect future cooperation to depend on their own behavior, then short-term defection would carry the larger long-term cost. Hence, incentives for cooperation strengthen even if the underlying preferences of the players remain unchanged.

Table 6
Illustrative Repeated Prisoner's Dilemma with Conditional Retaliation

Strategy	Round 1	Round 2	Round 3	Net Payoff
A and B cooperate throughout	(3, 3)	(3, 3)	(3, 3)	(9, 9)
A defects first; B responds by defecting thereafter	(4, 1)	(2, 2)	(2, 2)	(8, 5)

Note. Payoffs use the same hypothetical rankings as the original Prisoner's Dilemma. Player A can gain an immediate advantage by defecting in Round 1, receiving 4 instead of 3. However, if that defection causes player B to stop their cooperation in later rounds, A's total payoff falls to 8, compared with 9 under continued mutual cooperation. The model aims to illustrate how repeated interaction can make short-term defection less attractive by attaching future consequences. Author-created illustrative model based on the standard Prisoner's Dilemma framework.

We can additionally influence the game by altering social payoffs instead of merely economic ones. Human beings never truly evaluate outcomes solely according to money and financial gain. Reputation, societal approval / acceptance, group identity, personal values, ego, and social norms are all factors that contribute towards perceived utility. Cooperation that carries heavy economic costs may yet become desirable if it can also produce a substantial positive reputational benefit or if defection violates an already established social norm. In retrospect, social expectations effectively manifest themselves as part of the payoff matrix.

Conclusion

In conclusion, individual rationality alone cannot guarantee human progress. People make decisions according to their own preferences, experiences, uncertainties, goals, ideals, and egos. Yet, the accumulation of these decisions can produce consequences entirely contradictory to what any one person could have intended. Decision theory helps us explain how individuals attempt to choose decisions rationally, while heuristics, cognitive biases, and framing demonstrate why these decisions often depart from rationality. Simultaneously, paternalism further complicates attempts to correct said behaviors by asking how far society should interfere with individual autonomy in the pursuit of collective good.

The two cases of vaccination hesitancy and climate change demonstrate this conflict rather clearly. In both cases, collective welfare depends heavily upon widespread cooperation,

yet individuals possess incentives to avoid bearing some of the costs of that cooperation themselves. Vaccine hesitancy encourages individuals to benefit from the protection created by others, whilst climate change allows individuals to receive immediate benefits from environmentally costly behavior while also distributing much of the resulting harm across society and into the future.

Ultimately, game theory suggests to us that collective irrationality is not exactly solved by demanding that individuals stop pursuing their own interests. A society that's composed of self-interested individuals can still produce cooperative outcomes whenever its institutions structure incentives properly. Therefore, the best solution isn't suppression of self-interest, but rather the alignment of individual and collective rationality. By altering rewards, costs, expectations, and long-term consequences, society has a chance of transforming cooperation from an act of sacrifice into an act of rational self-interest. And is there some form of extent to this? Just how far should society reshape individual incentives without undermining individual freedom? Progress should depend upon finding an equilibrium that lies between autonomy and collective responsibility, enabling the very same destructive self-interest within all of us to become a force for our collective progress.

Works Cited

Betsch, Cornelia, Robert Böhm, and Lars Korn. “Inviting Free-Riders or Appealing to Prosocial Behavior? Game-Theoretical Reflections on Communicating Herd Immunity in Vaccine Advocacy.” Health Psychology, vol. 32, no. 9, 2013, pp. 978–985. https://doi.org/10.1037/a0031590.

Centers for Disease Control and Prevention. “History of Measles.” CDC, 29 Apr. 2026, https://www.cdc.gov/measles/about/history.html. Accessed 30 Aug. 2026.

Einhorn, Hillel J., and Robin M. Hogarth. “Behavioral Decision Theory: Processes of Judgment and Choice.” Journal of Accounting Research, vol. 19, no. 1, 1981, pp. 1–31. JSTOR, https://doi.org/10.2307/2490959.

Fine, Paul, Ken Eames, and David L. Heymann. “‘Herd Immunity’: A Rough Guide.” Clinical Infectious Diseases, vol. 52, no. 7, 2011, pp. 911–916. https://doi.org/10.1093/cid/cir007.

Gališanka, Andrius. “Just Society as a Fair Game: John Rawls and Game Theory in the 1950s.” Journal of the History of Ideas, vol. 78, no. 2, 2017, pp. 299–308. https://doi.org/10.1353/jhi.2017.0017.

“Game-Theory Models for Common-Resource, Public-Goods, and Prisoner’s Dilemma Incentive Structures” follow-up prompts. ChatGPT, model GPT-5.6 Sol, OpenAI, 30 Aug. 2026, https://chatgpt.com/.

Hart, Sergiu. “Adaptive Heuristics.” Econometrica, vol. 73, no. 5, Sept. 2005, pp. 1401–30. https://doi.org/10.1111/j.1468-0262.2005.00625.x.

Intergovernmental Panel on Climate Change. Climate Change 2023: Synthesis Report. Edited by Hoesung Lee and José Romero, IPCC, 2023. https://doi.org/10.59327/IPCC/AR6-9789291691647.

Lim, Wooyoung, and Pengfei Zhang. “Herd Immunity and a Vaccination Game: An Experimental Study.” PLOS ONE, vol. 15, no. 5, 2020, article e0232652. https://doi.org/10.1371/journal.pone.0232652.

North, D. Warner. “A Tutorial Introduction to Decision Theory.” IEEE Transactions on Systems Science and Cybernetics, vol. 4, no. 3, Sept. 1968, pp. 200–10. https://doi.org/10.1109/TSSC.1968.300114.

Swedberg, Richard. “Sociology and Game Theory: Contemporary and Historical Perspectives.” Theory and Society, vol. 30, no. 3, June 2001, pp. 301–35. JSTOR, https://www.jstor.org/stable/657964.

Trout, J. D. “Paternalism and Cognitive Bias.” Law and Philosophy, vol. 24, no. 4, July 2005, pp. 393–434. https://doi.org/10.1007/s10982-004-8197-3.

World Health Organization. “Poliomyelitis.” World Health Organization, 2 Apr. 2025, https://www.who.int/news-room/fact-sheets/detail/poliomyelitis. Accessed 30 Aug. 2026.

World Health Organization. “Vaccines and Immunization: Vaccine Safety.” World Health Organization, 23 Sept. 2025, https://www.who.int/news-room/questions-and-answers/item/vaccines-and-immunization-vaccine-safety. Accessed 30 Aug. 2026.

Greenwashing: A Review of the Linguistic Strategies Corporate Use By Lauren Kim

Abstract

This study examines how linguistic strategies can be used for corporate greenwashing. By investigating four linguistic strategies, such as the use of vague and ambiguous lexical choices, hedging through modal verbs, emotive and virtue-signaling language and framing, and using boilerplate repetition, this study supports the argument that greenwashing is not merely a marketing device, but a strategy that uses linguistic tools and reporting structures. This study offers insights for investors, corporate managers, regulators, and consumers to pay attention to corporate managers' word choice in reports and public disclosures with a critical mindset. This study reminds us that recognizing these linguistic strategies is an initial step toward demanding transparent corporate disclosures and proper responsibility for their environmental impacts.

Keywords Linguistic strategy; lexical choice; greenwashing

Introduction

In 1986, the term 'greenwashing' was coined by American ecologist Jay Westerveld from his visit to Fiji. He found two things that caused controversy at the time while staying at the hotel. Hotel management asked guests to reuse the towels to reduce the amount of detergent used. At the same time, the hotel's expansion may have done ecological harm (Cherry, 2013; Pizzetti et al., 2021; Poole, 2026). This anecdote illustrates what greenwashing really means: using rhetoric about the environment to mask hidden practices that make the environment worse. Over decades, as public awareness of anthropogenic climate change has grown, greenwashing has evolved from basic marketing into a sophisticated communication practice employed by major global corporations.

The purpose of this study is to investigate what linguistic strategies might be used in corporate greenwashing. For the scope of this paper, linguistic strategies are defined by the intentional choice of vocabulary, modality, rhetoric, and disclosure structure. Much business and economic academic research has examined why firms engage in greenwashing and its financial and economic consequences, but more linguistic and discourse-analytic studies are needed to investigate how they use linguistic strategies. A systematic linguistic and analytic approach is essential to understand how corporations construct environmental disclosures to manipulate stakeholder perceptions. By identifying the specific grammatical, lexical, and rhetorical choices in corporate writing, this study examines how language functions as a tool of impression management, and how four linguistic strategies can be used by companies as powerful devices of impression management for greenwashing.

The remainder of this study is organized as follows: the next section presents four linguistic strategies that firms can use. The last section concludes and discusses some limitations.

1. The Use of Vague and Ambiguous Lexical Choices

In business writing, vagueness may occur when a word has no clear limits. Ambiguity is found in a statement that can have more than one interpretation. Both vagueness and ambiguity can make claims harder to evaluate, reducing criticism and allowing companies to maintain a positive environmental reputation (Bloomfield, 2002; Poole, 2026). Prior literature documents that the management obfuscation hypothesis expects that managers have advantages to obfuscate information when their companies expect poor performance or underachievement. According to Bloomfield (2002), managers may use ambiguous linguistic choices to obfuscate information about firms' uncertain or poor future performance to postpone market reaction to bad news. Rooted in the management obfuscation hypothesis and the incomplete revelation hypothesis (Bloomfield, 2002), managers use vague terminology to increase information processing costs for analysts and the public.

In an analysis of John Browne (the CEO of BP, formerly British Petroleum)'s speech at Stanford University in 1997, Poole (2026) states that his speech is an instance of greenwashing through vague language and presenting possibilities rather than certain promises. His speech promises to *"increase the level of support we give to the continuing scientific work which is necessary,"* but there are no details given about the financial support level nor how much the support will be increased. Firms could engage in vague lexical choices to reduce immediate scrutiny while preserving future operational benefits.

2. Hedging Through Modal Verbs and "We believe" Phrases

Firms strategically employ modal verbs (such as may, might, will, would, can, could, etc.) to hedge environmental disclosures, thereby mitigating potential criticism and allegations of greenwashing. Poole (2026) states that the use of modal verbs and lexical choice is also a tool of hedging, which is shown in many greenwashing texts. They show an advertisement was placed in the New Statesman magazine by Luton Rising, a company that owns London Luton Airport, as below.

"Our unique Green Controlled Growth framework will introduce limits for the airport's noise, carbon, air quality and road traffic impacts. These would be legally binding and independently monitored."

The paragraph starts with the word *will*, indicating certainty, or a strong intention. The second sentence uses *would,* which makes the statement less certain and suggests the effectiveness of legal accountability and monitoring processes is not guaranteed. *Will* is a modal verb that presents a future event as expected (Merriam-Webster, n.d.). In contrast, *would* expresses conditionality, hypothetically, or greater distance from the proposition (Merriam-Webster, 2025). Critical readers may raise concern about the London Luton Airport expansion project because the shift from *will* to *would* is significant. The advertisement makes the environmental framework sound certain while using a weakened form when describing its accountability mechanisms.

Similarly, the paragraph below indicates a hedging strategy. The author uses the phrase “some of the far-reaching commitments” instead of “the most far-reaching commitments,” which indicates a hedging strategy. The use of 'some of the' indicates a clear example of hedging, as Lakoff (1973) argues that hedges are 'words whose job is to make things fuzzier'.

"Our Sustainability and Net Zero Strategies and our Green Controlled Growth approach to expansion, make, we believe, some of the most far-reaching commitments to minimizing environmental impact ever put forward by a UK airport"

By inserting the phrase “we believe,” a phrase which helps the company deal with the conflict between public demands for bold sustainability goals and financial risks of strict legal responsibility. Truly, the statement reflects the corporation's belief, but not an actual fact. Corporations may use hedging when future outcomes are uncertain. The greenwashing concern occurs when hedging is combined with ambitious claims that lack specific evidence.

3. Using Emotive and Virtue-Signaling Language and Framing

Virtue signaling can be understood as communication that displays moral values or socially desirable commitments. Such language helps corporations build a positive moral identity. The rhetorical emphasis may shift from measurable environmental outcomes to the corporation's apparent values and intentions. Framing refers to the way language selects and highlights aspects of an issue. Framing does not deny climate change. It emphasizes different causes, duties, and solutions. Framing influences what parts of a topic get noticed. It can also influence who appears more responsible for causing or solving the problem (Feeny et. al 2025). Framing is highly relevant to greenwashing because a corporation can redirect attention away from its own environmental responsibility by emphasizing another person's role. For example, the concept of an individual carbon footprint makes climate change through personal lifestyle. This makes the individual consumer an environmental actor. This focus on personal behavior can shift attention away from systemic and corporate sources of emissions.

Research on fossil fuel communication has shown that consumer demand and individual responsibility are overemphasized by corporations to construct a ‘demand-as-blame' narrative, labeling climate change as an issue of personal choice rather than fossil fuel production. (Supran & Oreskes, 2021). For example, in 2005, BP (formerly British Petroleum) hired a PR company to popularize the concept of individualized “carbon footprint" (Schendler, 2021). This rhetorical device aimed at attributing global carbon emissions to consumers. Their goal seemed to be achieved, but they cannot avoid the fact that their operation could cause a global catastrophe. Another example is the "Keep America Beautiful” campaign, which is best known as "The Crying Indian" commercial. This campaign supports packaging trade organizations by preventing the start of more rigorous recycling acts. This campaign, rather than attributing environmental degradation to the packaging industry, displaces accountability onto the public for illegal littering. Such examples can be viewed as a type of greenwashing that uses the virtue-signaling language of philanthropy to shift corporate environmental responsibility to customers and other stakeholders.

Emotive and virtue-signaling language is chosen by corporate managers because corporations don’t simply want to communicate information. Environmental communication may emphasize concepts like responsibility, leadership, and caring about communities. Such language creates positive associations with corporations and environmental responsibilities. Virtue signaling language is expressing moral values or beliefs to enhance one's social standing, rather than from genuine conviction. That means that companies can benefit from environmental language without providing evidence of environmental action.

4. Using Boilerplate Repetition

Boilerplate refers to a standard, reusable piece of context used in various types of content, including articles, press releases, and contracts (Mailchimp, 2023). It can serve legitimate purposes, but excessive boilerplate can reduce the informational value of corporate reporting. This happens when positive sustainability statements are repeated while company evidence or negative environmental information is limited. Therefore, it relates to greenwashing because it concerns the informativeness and accuracy of corporate environmental communication.

Firms strategically employ boilerplate repetition to comply with disclosure regulations while minimizing the revelation of negative material information. Lang and Stice-Lawrence (2015) state that regulators and standard setters identify boilerplate in financial reporting as problematic because standard setters such as IASB and FASB expect annual reports containing more boilerplate to be less informative. Hans Hoogervorst, chairman of the IASB, has stated that boilerplate in financial reporting is one of the primary concerns for standard setters, and it may even provide opportunities to hide information, reducing the informativeness of financial information (Hoogervorst, 2013). Measuring the readability of 10-K filings with the number of words and phrases, Li (2008) argues that increased textual noise and repetitive boilerplate writing may confuse readers with underachieved financial performance or hinder investors' ability to assess true environmental impacts of firms' operations.

A recent study by Kılınç et al. (2026) argues that companies strategically deploy boilerplate to emphasize positive initiatives and understate fewer positive actions when reporting sustainability information. Furthermore, Roberts et al. (2025) document that boilerplate such as indications of repetition, ambiguity, and weak information content can be used to hide inaction and deceive readers' perceptions. Boilerplate language is not inherently deceptive because standardized disclosure can serve genuine regulatory and reporting purposes. However, in sustainability reporting, boilerplate becomes relevant to greenwashing when generic environmental commitments substitute for measurable and company-specific information (Kilinc et al, 2016).

Conclusion

This study states that greenwashing is not merely a marketing device, but a strategy that uses linguistic tools and reporting structures. These strategies operate at different levels of corporate communication. Through ambiguous definitions, modal hedging, emotional/moral

imagery, and boilerplate, corporations present a picture of environmental responsibility but might be safe from measurable responsibilities. It is imperative for investors, corporate managers, regulators, and everyday consumers to accordingly pay attention to corporate managers' word choice in their reports or media disclosures with a skeptical mindset. This study highlights that greenwashing frequently occurs unnoticed in media advertisements, corporate reports or news, and social media content that provides information about products and services we consume. Recognizing these linguistic strategies is a critical step toward asking for transparent corporate disclosures and proper responsibility for their environmental influences.

Limitation

This study does not measure the environmental performance of corporations. Rather, it discusses the linguistic and reporting practices companies use. The four linguistic strategies discussed do not comprise all the strategies. Different industries have different forms of environmental communication. Thus, the results presented should be interpreted as linguistic indicators associated with greenwashing rather than definitive evidence that greenwashing has occurred.

Works Cited

Bloomfield, R. J. (2002). The “incomplete revelation hypothesis” and financial reporting. Accounting Horizons, 16, 233–243.

Cherry, M. A. (2013). The law and economics of corporate social responsibility and greenwashing. UC Davis Business Law Journal, 14, 281–299.

Feeney, M., Ormiston, J., & Gijselaers, W. (2025). Framing collective moral responsibility for climate change: A longitudinal frame analysis of energy company climate reporting. Journal of Business Ethics, 198, 485–508. https://doi.org/10.1007/s10551-024-05801-0

Hoogervorst, H. (2013, June 27). Breaking the boilerplate [Conference remarks]. IFRS Conference, Amsterdam, Netherlands.

Kılınç, Y., İnce, M. R., & Badem, A. C. (2026). A multi-dimensional textual framework for detecting greenwashing in sustainability reporting. Discover Sustainability, 7, Article 482.

Lakoff, R. (1973). Language and woman's place. Language in Society, 2(1), 45–79.

Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. Journal of Accounting and Economics, 60(2–3), 110–135. https://doi.org/10.1016/j.jacceco.2015.09.002

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45(2–3), 221–247.

Mailchimp. (2023). Crafting the perfect boilerplate. https://mailchimp.com/resources/boilerplate/

Merriam-Webster. (n.d.). Will. In Merriam-Webster.com dictionary. Retrieved August 25, 2026, from https://www.merriam-webster.com/dictionary/will

Merriam-Webster. (2025). Would. In Merriam-Webster.com dictionary. Retrieved from https://www.merriam-webster.com/dictionary/would

Pizzetti, M., Gatti, L., & Seele, P. (2021). Firms talk, suppliers walk: Analyzing the locus of greenwashing in the blame game and introducing “vicarious greenwashing.” Journal of Business Ethics, 170(1), 21–38.

Poole, B. (2026). Linguistic features associated with greenwashing. Cogent Arts & Humanities, Article 2653345. https://doi.org/10.1080/23311983.2026.2653345

Roberts, L., Wang, Y., & Elamer, A. A. (2025). Do vice chancellors' career horizon matter for university sustainability performance? The moderating role of soft information. International Journal of Finance & Economics, 30(3), 2964–2980.

Schendler, A. (2021, August 31). Worrying about your carbon footprint is exactly what big oil wants you to do. The New York Times.

Supran, G., & Oreskes, N. (2021). Rhetoric and frame analysis of ExxonMobil's climate change communications. One Earth, 4(5), 696–719. https://doi.org/10.1016/j.oneear.2021.04.014

Images of Violence: An Analysis of Cartel Imagery in Media and Its Impact on Mexican Society By Alston Guo

Abstract

Mexican Cartels are increasingly using mainstream and digital media to influence public opinion and to increase the presence they have in society. What are the effects of the representation of Mexican cartels in mainstream and digital media? This literature review analyzes research on the effects of the representation of Mexican Cartels in mainstream and digital media. Research clusters around the effects on government legitimacy, security, and youth recruitment. Overall, the findings demonstrate that the media provides a means for cartels to influence the public's perception of them and solidifies their overall influence over Mexican society.

Evolution of Cartel Communication

What are the effects of the representation of Mexican cartels in mainstream and digital media? In Mexico today the predominant tools used by criminal organizations beyond displays of violence are the deliberate creation of narratives and the strategic development of discourse. These narratives and discourses are the products of decades of evolution and were created specifically to meet the needs of Mexican cartel operations. They are being used today to control how people think about criminal organizations and how information about cartel operations is shared (Sullivan 2012). Thus, the way in which Mexican cartels communicate is reflected in a broader pattern in mainstream and digital media, which influences Mexican civilians. Accordingly, a growing number of scholars have examined the extent to which the media affects government legitimacy, public security, and youth recruitment in Mexico.

Mexican drug cartels began in earnest during the early part of the 20th Century. The decline of major Colombian cartels allowed for the eventual rise of other major Mexican cartels. The concentration of power under individuals like Miguel Ángel Félix Gallardo helped establish Mexico as a major player in the global drug trade (Beittel 2012). Initially, early cartels communicated through banners placed throughout cities, large-scale public displays of violence, and through word-of-mouth intimidation. However, in the 2010s, these cartels began utilizing digital platforms to facilitate communication between cartel associates, to spread propaganda, and to issue threats to potential adversaries or the Mexican State. Social media has become an outlet for coded communication and targeted recruitment. Major players in the Mexican cartel world, such as CJNG (Jalisco New Generation Cartel), have utilized various forms of social media to recruit new members, to advertise illegal drug sales, to distribute false information, and to project an image of financial prosperity and impunity (International Crisis Group 2024).

Legitimacy of Power and Governments

The media has come to play an increasing role in defining power and legitimacy. Mexican cartels use the media to develop and manage their own narratives, therefore reducing

the legitimacy of the government. By using aggressive methods to intimidate journalists, purchasing advertising space in newspapers, compelling print media outlets to publish their propaganda, and issuing pseudo-public service announcements, cartels aim to portray themselves as more dependable, competent, and efficient than government institutions. Cartels produce media that focus on themes of order, security, and social services, especially in regions with a weak government presence. Therefore, by framing cartels as more organized and responsive than official government agencies, cartels can expand their influence. Within this context, the media has evolved as a cost-effective vehicle for cartels to assert dominance and exert influence.

Although some scholars contend that direct experience with violence would be the main cause of fear and distrust, others argue that media portrayals contribute to fear and mistrust. Vilalta Perdomo (2013) conducted a peer-reviewed study examining the effects of the media on fear and mistrust in Mexico. Her results indicated that many Mexicans have experienced violence firsthand. Thus, their fear may be influenced less by media portrayals and more directly by their immediate personal risk for harm. Carreras (2013) similarly finds that direct victimization and higher perceptions of violence can reduce support for political institutions not just in Mexico but in broader Latin America.

While Vilalta Perdomo (2013) concludes her study by stating that she could not confirm the extent to which media influences narratives of fear, she does not entirely reject the possibility that media affects such narratives. Regardless, while Vilalta Perdomo's argument highlights the significance of living experiences, it distinguishes between experiences with violence and media. Taken together, these studies suggest that direct experiences with violence and the way media frames narratives may interact in shaping the perception of how citizens interpret security and therefore, media may provide an additional perspective through which experiences of violence are perceived. That said, in addition to shaping how violence is interpreted, the media portrays existing conditions and creates narratives that define Mexico's political landscape, thereby supporting the assertion that it is a critical factor in shaping public perception and undermining democratic institutions. Furthermore, Mexico's recent electoral participation data support this contention. Specifically, Mexico's last judicial election had a voter turnout of just 13%, reflecting broader worries about political participation and institutional legitimacy (Graham 2025).

Security

Existing research finds that Cartel violence portrayed in mainstream media significantly alters civilians' perceptions of what constitutes danger. However, the media does so by portraying a disproportionate amount of instability in Mexico. Additionally, entertainment media significantly affects how civilians perceive danger. Many television programs and films portray Mexican cartels as organized, technologically advanced, and capable of challenging state authorities. Furthermore, these programs frequently exaggerate violence visually and narratively to portray Mexico as inherently violent, causing many viewers to internalize these representations despite their fictional nature. This dynamic is explored in Rodríguez-Blanco and Mastrogiovanni (2018), who state that sensationalist reporting focusing primarily on

kidnappings, assassinations, and armed confrontations is designed to maximize viewer engagement. These findings suggest that media representations may perhaps contribute to a disconnect between perceived and experienced vulnerability. For example, surveys conducted by the Pew Research Center consistently show that Mexicans rank violence, corruption, and cartel-related problems as among Mexico's top problems, regardless of whether the survey respondents live in areas known for having low crime rates (Heimlich 2011). To create the impression of omnipresence, cartels regularly post videos documenting targeted killings or attacks on public figures to project power. Rather than providing factual information regarding crime in Mexico, Rodríguez-Blanco and Mastrogiovanni's (2018) analysis suggests that sensationalized media reporting exacerbates civilians' fears of insecurity.

Rodríguez-Blanco and Mastrogiovanni (2018) further analyze how media reports often employ stereotypes based on visual cues (e.g., yellow filters) and often use these cues to magnify violence and simplify images of Mexico. These depictions have real-world implications: sensationalized reporting focused on kidnappings, shootouts and corruption can discourage foreign investors from investing in Mexico, as well as discourage tourists from travelling there. Fears, Lies, and Lucre, in a publication by the International Crisis Group (2024) found that local businesses operating in areas with cartel activity regularly report extortionate demands from criminals acting on behalf of cartels. Furthermore, according to Morris (2012), corruption enables the existence of criminal organizations while simultaneously reducing the Mexican state's capacity to combat cartels.

Moris (2012) finds that when civilians feel that governments cannot protect them from harm or enforce democratic principles, civic participation typically decreases. The decrease in civic participation limits governments' abilities to respond to threats posed by criminal organizations. Therefore, media narratives not only represent current conditions but also help create Mexico's political climate. In doing so, the media helps create an environment conducive to continued criminal dominance and increased social inequality in Mexico. Moris (2012) further argues that corruption increases the amount of drug trafficking while weakening the state's capacity to confront criminal organizations.

Youth Recruitment

Existing research examines how glorified presentations of Mexican cartels through social media illustrate how young adults' value systems are formed in Mexico and facilitate the recruitment process of youth into cartel structures. Social media creates pathways for youth involved in cartel recruitment to transition into more serious criminal behaviour. The promotion of narcoculture through music, images, and short-form video formats, delivered via digital media, presents positive portrayals of wealth, power, and prestige associated with participation in cartel organizations. Youth recruited into cartel structures perceive membership as a viable route to upward social mobility, as demonstrated by a 2023 study by Tyler Biscontini. This phenomenon is enhanced in impoverished communities due to limited job opportunities. Cartels take advantage of this structural disadvantage by recruiting through digital platforms such as

WhatsApp and Telegram, often using emotional manipulation through messages presenting cartel members as protectors/ providers. In summary, social media serves not only as a platform promoting cartel culture but also as a means of recruiting youth into criminal behaviours, conditioning them to participate, and financially incentivizing their involvement. Ultimately, all three factors suggest an increase in the overall strength and duration of Mexican cartels.

Gil (2024) states that poverty, unemployment, and limited opportunities for social mobility are primary causes for cartel recruitment. Advocates for this viewpoint suggest that economic hardship is the primary driver motivating youth participation in cartel activity. However, although poverty may render youth more susceptible to recruitment efforts by cartels, poverty alone does not account for why youth choose to participate in cartel activity. Instead, youth are attracted to online portrayals that emphasize the wealth, power, and prestige associated with cartel membership.

Conclusion

One conclusion reached from this paper is that Mexican cartels are able to leverage media as a tool for exerting influence upon society. Therefore, altering societal perceptions relative to government efficacy. Citizens exposed daily through social media to continuous images depicting violence are more likely to experience fear and less likely to engage civically. Moreover, consistent media coverage of corruption scandals, coupled with portrayals of cartel dominance, undermines citizens' faith in institutional authority, thereby reducing reporting of crimes to authorities.

This paper highlights the important premise that media can not only reflect cartel influence but also enhance its influence. However, this research is subject to several constraints, specifically the limited availability of platform algorithmic data and reliance on secondary sources. Despite these constraints, this paper highlights an urgent need for a comprehensive approach to counteract cartels' persuasive influence in digital environments through national programs that educate digitally literate citizens, localized approaches that support youth development, independent journalistic endeavours, and fundamental socio-economic reform to mitigate the broader impact of cartels. If left unchecked, cartel influence will evolve solely through continued usage of media as a conduit that sustains both criminal dominance and social inequality in Mexico.

Future research should focus on using quantitative data to better understand the impact of cartel-related media on Mexican society. There have been some previous studies that have used quantitative data to measure media exposure, public perceptions of insecurity, and patterns of online cartel recruitment. Future researchers can use large-scale surveys to examine whether there is an association between the amount of cartel-related media exposure civilians experience and a lack of confidence in governments, as well as perceptions of security and legitimacy. Furthermore, researchers can statistically analyze regions with varying levels of cartel activity and draw conclusions based on varying statistics. Lastly, increased access to algorithmic data, such as engagement rates, recommendations for cartel-related content, and the geographic

distribution of cartel-related content, would help researchers identify populations more likely to be exposed to this type of media. The combination of these types of quantitative evidence will provide a much clearer picture of how the media influences poverty, violence, and limited economic opportunity.

Works Cited

Beittel, J. S. (2012, June 8). Mexico's Drug Trafficking Organizations: Source and Scope of the Rising Violence. U.S. Department of Justice. https://www.justice.gov/file/267651/dl?inline=

Biscontini, T. (2023). Narcoculture in Mexico: Research starters: EBSCO research. EBSCO. https://www.ebsco.com/research-starters/law/narcoculture-mexico

Carreras, M. (2013). The Impact of Criminal Violence on Regime Legitimacy in Latin America. Latin American Research Review 48(3), 85-107. https://dx.doi.org/10.1353/lar.2013.0040.

Gil, R. Z. (2024). Mexican government failing to provide decent jobs for vulnerable youth – leaving the door open to cartel recruitment: Oxford Department of International Development. Oxford Department of International Development. https://www.qeh.ox.ac.uk/blog/mexican-government-failing-provide-decent-jobs-vulnerable-youth-leaving-door-open-cartel

Graham, T. (2025, June 2). Mexican president hails "complete success" after just 13% vote in judicial elections. The Guardian. https://www.theguardian.com/world/2025/jun/02/mexico-judicial-election-claudia-sheinbaum

Heimlich, R. (2011, August 31). Crime and drug cartels top concerns in Mexico. Pew Research Center. https://www.pewresearch.org/global/2011/08/31/crime-and-drug-cartels-top-concerns-in-mexico/

International Crisis Group. (2024). Fear, Lies and Lucre: How Criminal Groups Weaponize Social Media in Mexico. International Crisis Group. https://www.crisisgroup.org/latin-america-caribbean/mexico/b50-fear-lies-lucre-how-criminal-groups-weaponise-social-media-mexico

Morris, S. D. (2012). Corruption, Drug Trafficking, and Violence in Mexico. The Brown Journal of World Affairs, 18(2), 29–43.

Rodríguez-Blanco, S. and Mastrogiovanni, F. (2018). Hegemonic narratives of violence. organised crime and drug trafficking between journalism and television fiction. Anàlisi, (58), 89. https://doi.org/10.5565/rev/analisi.3098

Sullivan, J. P. (2012). Criminal insurgency: Narcocultura, social banditry, and information operations. Criminal Insurgency: Narcocultura, Social Banditry, and Information Operations | Small Wars Journal. https://archive.smallwarsjournal.com/jrnl/art/criminal-insurgency-narcocultura-social-banditry-and-information-operations

Vilalta Perdomo, C. J. (2013). Determinant Factors in the Perception of Crime-Related Insecurity in Mexico. https://doi.org/10.18235/0011448

Where the Music Takes You: Autobiographical Recall vs. Fictional Narrative Across Classical Subgenres By Derrick Lin

Abstract

Western classical music is frequently treated as a single genre in psychological research despite encompassing centuries of radically different compositional traditions. This paper examines how six major periods of Western classical music—Baroque, Classical, Romantic, Impressionist, Modernist, and Contemporary—may differently encourage autobiographical memory and fictional narrative during listening. This paper compares the compositional characteristics associated with each period and their potential cognitive consequences. Further, existing research suggests that familiarity provides strong cues for autobiographical recall, while unfamiliar, dynamic, or culturally narrative forms of music more readily facilitate imaginative scenes and fictional narratives. These findings complicate the conventional treatment of classical music as a single psychological category. This paper argues that music-evoked thought results from the interaction between musical structure and listeners' prior experience rather than genre alone. A more fine-grained understanding of these relationships may contribute to future research in music cognition, therapeutic applications, and narrative media composition.

I. Introduction

Music has a special ability to pull the listener to a multitude of imaginative responses. Most prominent among those responses are vivid personal memories—detailing past settings, relationships, and moments—and scenes that never existed at all: fictional narratives constructed with creativity. These two responses, the reliving of personal experience and the generation of fictional narrative, represent fundamentally different cognitive operations. One is anchored in the self and the past; the other is untethered and isolated from it. Yet both emerge spontaneously and reliably during music listening, often requiring no extensive effort on the listener's end. What determines which response a piece of music elicits remains an ongoing and open investigation in the psychology of music.

Historically, empirical literature has relied on broad genre classifications to assess musical stimulus responses, which leaves a critical gap in investigating the specific micro-level details that exist within each genre. Most notably, existing studies often group Western classical music as one uniform variable, overlooking the centuries of diverse and contrasting musical and historical backgrounds that exist in each period. This paper examines how structural and historical differences among Western classical subgenres may shape two major forms of thought: autobiographical recall and fictional narrative. Drawing on research in music cognition alongside comparative analysis of Baroque, Classical, Romantic, Impressionist, Modernist, and Contemporary music, it argues that musical structure, perceived energy, and listener familiarity jointly influence the kind of music-evoked thought to which music directs.

II. Historical and Musical Foundations

Classical music is not a monolith. The term "classical music" often encompasses several centuries of compositionally, philosophically, and emotionally distinct traditions, each crafted on differing societal views of what music is and what its intended effect on listeners should be. Within the genre of classical music exists an ecosystem of subgenres, all fundamentally unique in their sound and intentions. For the purposes of this review, the term 'classical music subgenre' refers to four historically and philosophically differing periods within the Western art music canon: Baroque, Classical, Romantic, Impressionism, Modernism and Contemporary. These periods represent the most meaningfully distinct changes throughout classical musical history in sound and philosophy and therefore are the most relevant to the theoretical framework at hand.

In order to understand the connection of classical subgenres and the existing research upon genre as a determinant for listener response, the broader musical theory behind these periods must be discussed.

II. 1. The Baroque Period (c. 1600-1750)

The Baroque period was shaped by a revival of ancient Greek and Roman thought; in particular, the belief that music held the power to shape human emotion (Baroque Music Website). Composers of this time period included prominent names such as Bach, Monteverdi, Handel, and Vivaldi. In all their works, this greater philosophy of the time was used as a framework to harness emotion in their compositions. Central to Baroque music was a clearer distinction between melody and harmony, alongside techniques such as ornamentation, polyphonic counterpoint, terraced dynamics, canons, and transformations (Grout and Palisca; Burkholder et al.).

Monteverdi formalized two contrasting approaches: one prioritized harmony and counterpoint, while the other emphasized expressing the meaning of the text and shaping the listener's emotional response. The latter became especially influential in opera and dramatic vocal music, illustrating how Baroque music combined formal complexity with expressive intent. Baroque music, as an art form, was philosophically formed to invoke a deep emotional response from listeners while simultaneously using deliberate structure and form to enhance the narrative nature of the art (Burkholder et al.).

II. 2. The Classical Era (c. 1750-1820)

The Classical period is often coined as the "Age of Reason" or "Age of Enlightenment", both because it exists in the broader historical context of the Enlightenment and its stark contrasts to the previous Baroque period in terms of philosophy and musical structure (Boumpani and Carteret). This era was positioned under the same doctrine of the Enlightenment ideals: logic, clarity, and formal structure. The ornate complexity of Baroque musical architecture and composition was widely abandoned in favor of a more restrained, functional, and proportional way of music (Rosen).

Composers of this time period, consisting of Mozart, Haydn, Clementi, Schubert, and later, Beethoven, all adapted precise and replicable forms to organize music with the same systematicity that Enlightenment thinkers applied to other arts and sciences (Boumpani and Carteret). These musical aims were furthered by compositional techniques such as texture, tempo, dynamics, melody, and harmony.

II. 3. The Romantic Century (c. 1800-1900)

The Romantic period represented a conscious reaction against Classical restraint (Taruskin). Global developments such as the Industrial Revolution quickly spurred the philosophy of the time to develop into one that valued emotion, imagination, and individuality above formal prescription, as seen through various art forms, including music. Crucially, Romantic period music saw a change in the philosophical status of music. Previously, music was seen as inferior due to its lack of words and structure compared to literature, poetry, sculpture, and other art forms. However, the Romantics reframed this absence to synthesize music's greatest virtue—the capacity to "express what words could not express, thus transcending the material for something more ideal and spiritual" (Kluball and Kramer 45).

This philosophical rebranding gave way to program music—instrumental works tied explicitly to narrative, nature, literature, or visual art—as well as a new culture of heightened emotional directness in composition (Taruskin). Key composers of the time period consisted of Liszt, Rachmaninoff, Schubert, Chopin, and Tchaikovsky. These composers, inspired by various world and cultural events, would go on to write some of the most expressive pieces of music marked by a series of new compositional techniques that expanded on previous periods. Harmonic language became increasingly complex, dissonant, and chromatic, with frequent modulations to establish disorientation in listeners (Burkholder et al.).

II. 4. Impressionism (c. 1890-1920)

As the transitional bridge between late Romantic emotionalism and twentieth-century radicalism, the Impressionist movement captured moods, sensations, and atmospheres in a flowing artistic style (Pasler). Emerging primarily in France as a parallel to the visual arts movement of figures like Monet and literary movement figures like Mallarmé, Impressionist music rejected both the strict, rigid structures of Classical-Baroque music and the heavy dramatic weight of Romantic music (Palmer). Instead, the goals of Impressionist composers such as Debussy and Ravel were to allow the listener's imagination to drift through a variety of musical textures. Philosophically, the movement treated sound not just as a catalyst to drive emotional stories, but as an independent physical phenomenon.

II. 5. The Modernism Shift (c. 1900-1945)

The turn of the twentieth century marked the radical departure from the established traditions and structure of Western classical music. Perhaps most notable and famous of the modern musical movements is the Modernism movement led by composers such as Schoenberg,

Stravinsky, and Berg. This era was defined by the philosophical rebellion against the emotional and predictable structures of not only the Romantic period music but also the fundamental understandings of music built on the previous centuries. Modernist composers operated under the belief that the traditional tonal system, which dictated the rules of harmony that had governed classical music, had been exhausted and pointless (Schoenberg).

To achieve such a philosophical shift, Modernist composers replaced the standardized music theory with entirely new elements. The most significant breakthrough of this period was the introduction of atonality and serialism, for example, with Schoenberg's twelve-tone technique (Schoenberg). These structures and concepts were often utilized in Modernist music, meaning that there was no "home key" or tonic note, creating instability in the music from the beginning of the piece. Without a central key, the music exists in a state of constant harmonic suspension, allowing the composer to deny the listener a natural sense of arrival or conclusion found in earlier musical periods (Morgan).

Beyond the breakaway from harmonic tradition, the internal architecture of Modernist music relies on experimental dissonance and rhythmic asymmetry (Ross). The beauty of Modernist music lies in the fact that it was unpredictable to the general audience (Morgan). By combining structural atonality with abnormal and aggressive musical techniques, the Modernist era crafted a highly complex sonic adventure that subverted the human ear's natural expectation and instinct for pattern, melody, and structure (Ross).

II. 6. The Contemporary Postmodernism (c. 1945-Present)

The final period of music to be examined in this study is the Contemporary period, also known as Postmodernism. This period directly follows the wake of World War II and the rigid and radical intellectual musical rebellion of Modernist music. Postmodernism came to represent an era of total stylistic liberation, challenging the fundamental definition of what constitutes "music" even further and continuing to reject a linear historical trajectory (Metzer). The surrounding philosophies of the late twentieth and early twenty-first centuries were shaped by rapid globalization, technological revolutions, and the collapse of traditional cultural hierarchies. As the world became more interconnected than ever before, the historical boundary between performer, composer, and audience blurred, resulting in a fragmented musical ecosystem where radical Avant-Garde music could coexist with other styles of Contemporary music (Ross). Consequently, the music theory and structure of contemporary music fractured into a variety of wildly differing techniques that challenged traditional listener expectations.

By blending itself seamlessly with mass media, the Contemporary subgenre has transformed the modern definition of the classical canon into a creative medium spanning art forms that take on numerous compositions, all aimed at evoking difficult and complex emotions and thoughts about the everyday world.

III. Autobiographical Recall vs. Fictional Narrative

III. 1. Music-Evoked Narrative and Imagination

The relationship between music and creative imagination has become one of the more substantially researched areas within the psychology of music in recent decades. The foundation for this claim begins with Margulis, whose study provided the first systematic evidence that listeners could construct coherent stories in response to wordless instrumental music (qtd. in Jakubowski et al.). Margulis found that even in the absence of clear lyrical directions, listeners were still able to produce meaningful and vivid narratives, showing music's innate ability to evoke strong creative imaginations. Strikingly, she found that these spontaneous imaginings are not purely idiosyncratic; instead, when listeners had similar cultural backgrounds, an "intersubjectivity" was displayed. This created trends where listeners from the same cultures envisioned similar plots, character arcs, and emotional trajectories when listening to the exact same musical excerpts. Conversely, the opposite is found to be true; using advanced natural language processing (NLP) tools to analyze free-response descriptions, cross-cultural examinations demonstrated that the narrative consensus disappears when listeners do not share a cultural background (Margulis et al.). This means that while music has been shown to prompt people to imagine stories, a striking nuance in cultural background exists as a shared lens to shape stories similarly and predictably to peers.

As the music rolls, a complex cross-sensory translation occurs within the listener's mind, where the brain actively bridges the gap between auditory perception and visual-narrative synthesis. Analysis from Hashim et al. suggests that storytelling behavior is closely linked to visual imagery. They found that an astonishing 94.6% of adults reported they imagined vivid visual imagery while listening. Most of the images took the form of continuous, evolving stories rather than static snapshots. This study suggests that the mind may continuously and consistently translate dynamic auditory cues and changes, such as tempo, rhythm, timbre, pitch, and harmony, into physical imagery. For instance, a prolonged crescendo in combination with accelerando in the auditory cue could be cognitively interpreted by the brain as an escalation in situational tension. Although specific details of narration may vary in each listener's mind, drastic changes to the music most often elicit a natural response in imagination. The brain next maps these thoughts into a mental movie scene, translating the accompanying auditory stimuli to coherent narratives, characters, scenery, and interaction (Hashim et al.).

To understand how this creative-world-building is shaped in the mind, psychology research has pointed towards a phenomenon known as "cognitive decoupling" (Smallwood and Schooler). When an individual is listening to music, their attention shifts away from "focal focus" from the external environment to a state of internal stimulus-driven thought. In this state, the brain's Default Mode Network (DMN) – the neural framework responsible for internally oriented cognition, including mental simulation and perspective-taking – becomes heavily involved and activated (Taruffi et al.). Studies confirm that the DMN is stimulated positively depending on the emotional valence of the music, signaling that the mind utilizes the given auditory stimuli as a direct source for neurological communication, self-reflected thought, and imagination (Taruffi et al.).

III. 2. Music-Evoked Autobiographical Memory

While music has been shown to be a powerful trigger for creative imaginations, a parallel and emerging body of research showcases music's ability to produce a second distinct category of listener response: the spontaneous and vivid recall of autobiographical memories. Music-evoked autobiographical memories, commonly referred to in literature as MEAMs, are defined as the recollection of personal experiences and events triggered after a musical stimulus. The phenomenon has been proven to be a universal experience, regardless of background in musical training, which has spurred empirical investigation (Janata et al.). These two dominant response types are not unrelated phenomena, but rather two outcomes of the underlying capacity music has to direct a listener's inner world beyond the present moment.

The foundational contribution to this field stems from Belfi, Karlan, and Tranel, whose study tested the assumption that music is a uniquely powerful memory cue. The study compared memories evoked by music to those by famous faces and showed that music not only effectively evokes recall but also contributes to more vivid memories. This finding was significant because it signaled the shift from intuitive claims about music and memory towards grounded evidence that the art form could be used as a potent retrieval cue.

Building upon this foundation, Jakubowski and Ghosh examined how MEAMs operate outside of controller laboratory conditions, tracking MEAMs across participants' everyday lives rather than in experimentally designed settings. They found that these memories were triggered on average once a day and were typically rated highly and involuntarily. This study is significant due to its evidence surrounding the involuntary nature of these memory retrievals. The distinction between involuntary and voluntary retrieval is theoretically important as it suggests that music bypasses the deliberate cognitive effort typically associated with memory recall. Autobiographical memories have been proven in the past to either be recalled voluntarily, requiring effortful search and retrieval, or involuntarily, which seem to spontaneously come to mind without any intention to retrieve (Berntsen). Involuntary memories are not simply a passive version of recall; they are more likely to result in higher impacts on mood, bodily reactions, speed, as well as vividness (Berntsen and Hall). Music, as a predominantly involuntary retrieval cue, therefore delivers memories with a specificity and emotional impact that cannot be easily replicated.

To fully understand how these involuntary memories are happening in relation to music, it is necessary to examine the underlying neural components and architecture. Neurological studies confirm that MEAMs rely on the activation of a specific, distinct network of brain regions in the medial prefrontal cortex (mPFC). Importantly, the mPFC is located in the DMN, which, as previously stated, is the core hub for cognitive processing during wakeful rest, mind-wandering, and spontaneous, stimulus-independent thought (Janata). When a personal experience is recalled, the mPFC works together with the hippocampus and the parahippocampal gyrus to reconstruct the contextual details of the memory (Belfi et al.). Additionally, the mPFC is structurally tied to the amygdala and the brain's main reward centers, which prompts uniquely vivid and saturated emotions alongside the retrieved memories. The tight neurological

connection between the mPFC, activated through the DMN, and networks associated with emotional and memory processing explains why musical stimuli can instantly trigger highly specific past events without requiring effortful and conscious cognitive searching.

IV. Musical Genres and their Relation to Music-Evoked Thought

While the literature clearly shows autobiographical memory and fictional narrative as two primary pathways of music-evoked thought, a critical question remains regarding what prompts a listener's mind down one track over the other. Recent cognitive psychology has aimed to tackle this matter by examining how the mind's internal landscape is shaped by musical stylistic differences. Rather than triggering random and unpredictable imaginations, different genres and musical styles actively bias the human brain towards different profiles of thought.

The most comprehensive investigation of this distinction is offered by van der Walle et al., who tracked music-evoked thought (MET) across 17 distinct musical genres. By categorizing listener responses into imagination baskets, including fictional narration and autobiographical recall, their text provides empirical proof that thought type can shift predictably depending on the specific elements of the auditory stimuli.

A central contribution by Van der Walle et al. is their data showing specific acoustic and cultural profiles prompting entirely different modes of mental processing. For example, genres that carry a high level of familiarity or are rooted in mainstream popular culture, such as pop, rock, and hip hop, act as strong cues for autobiographical memory. Because listeners typically have a lifetime of exposure to these conventional rhythmic and melodic structures, these genres trigger a significant spike in retrieval. Conversely, genres tied directly to media or abstract environments, such as film music, ambient music, and "classical music", tend to steer the mind away from personal history. These genres, oftentimes, exemplify dynamic forms of music that bias the cognitive pathway towards vivid creative storytelling and imaginings.

Beyond broad genre labels, the researchers noted that the track's independent features such as "contrast" and "arousal" (energy) serve as underlying cognitive mechanisms that subtly direct thought types (Jakubowski et al.). Musical pieces characterized by high structural contrast, meaning sharp and sudden shifts in musical texture, tend to hold the listener's attention at higher rates. This dynamic characterization of music forces the brain to actively process auditory changes, drastically reducing the amount of mundane everyday daydreams ("mind-wandering"). Shown by their data, the most prominent pathways the brain takes outside of general musical thoughts about the given stimuli (52.6% of trials) were media memories (23.6%), fictional imaginings (22.5%), and autobiographical memories (16.4%) (van der Walle et al.). This contrasts with the other thought types, which all sat at less than 10% of listener response. Interestingly, the study also discovered that high arousal rate alone, high energy, and fast tempo music that remains structurally the same actually results in an increase in thoughts about the everyday (van der Walle et al.). This suggests that energy monotony alone is not enough to spark the brain to build dynamic and deep-creative thought.

Finally, one important nuance when analysing METs lies in the listener's familiarity. Although the musical genre demonstrates a high level of correlation with which specific path the listener's thought type is skewed in, one's own personal background influences and potentially disrupts the baseline genre expectations. This complexity serves as a critical moderating lens through which the music's acoustic profile is filtered (van der Walle et al.). Mainstream modern genres such as pop, hip-hop, and rock are structurally built upon recognizable patterns that prompt high familiarity rates among contemporary listeners. This high rate of familiarity prompts higher rates of autobiographical recall as it anchors the cognitive pathways in real-world stimuli. Conversely, genres that are lower in familiarity for the general audience, such as classical music and complex electronic genres, strip away those personal cues. By removing the listener's cognitive familiarity with the musical stimuli, these tracks tend to evoke more creative and narrative responses (Margulis et al.). Therefore, investigating genre relation to METs must crucially also acknowledge the personal and individualistic relations of each listener's familiarity with those stated genres. For instance, participants more familiar with classical genres than pop genres do not demonstrate contrasting results of MET compared with participants who listen to modern genres on a daily basis (van der Walle et al.). The existing literature provides strong empirical evidence that genre in the context of a general audience shows predictable and consistent results in thought type; however, it is also important to note that this relation differs on an individualistic scale due to familiarity being a variable.

Altogether, the existing literature continues to strongly demonstrate music's unique ability to evoke complex and vivid imaginations within listeners. Most notably amongst these thought types lie fictional narratives and autobiographical memory, triggered by different aspects of music and cognitive functions. Although recent research has begun investigating how those two pathways are connected to each other and specific musical elements across broad genres, there remains a distinct shortage of research regarding what specific musical techniques prompt such responses. In order to address this gap, the proposed pilot study will investigate METs through six distinct Western classical subgenres: Baroque, Classical, Romantic, Impressionism, Modernism, and Contemporary. These subgenres, spanning across centuries of differing histories and philosophies, are fundamentally different from one another in sound and structure. All six subgenres, however, provide an ideal framework for investigating the specific details behind music's ability to generate METs due to the deep-rooted connection with music theory and the exceptionally extensive vocabulary of terminology available.

V. Proposal for Future Research

While foundational literature in music psychology has robustly explored broad genre categories (van der Walle et al.) as well as the existence of METs and MEAMs (Margulis, 2017; Belfi et al., 2016), more studies are required to bridge a crucial gap by narrowing the analytical lens to micro-movements in the classical music genre. Existing frameworks often categorize classical music under one umbrella, which ultimately fails to account for the complex nuances between each sub-genre, such as a Baroque fugue in contrast with a Romantic concerto. To

illustrate the differences among these periods, consider some representative works by composers including Bach, Telemann, Beethoven, Rachmaninoff, Ravel, Debussy, Shostakovich, Webern, Philip Glass, and Caroline Shaw.

Musical Era	Example 1	Example 2
Baroque	Bach — Harpsichord Concerto No. 1 in D minor BWV 1052: III	G. Ph. Telemann — Sonata in D major, TWV 44:1
Classical	Joseph Boulogne — Violin Concerto in A major, Op. 5 No. 2 (1st Movement)	Beethoven — Pastoral Op 28 Sonata MVT II
Romantic	Rachmaninoff — Prelude Op 32 No 12	Schumann — Piano Quartet / Brahms — 2nd Symphony MVT II
Impressionist	Ravel — Sonatine, M. 40: I. Modéré	Debussy — La Mer MVT II
Modern	Shostakovich — String Quartet No. 8 in C minor Op. 110: II (or 10th Symphony)	Webern — Three Little Pieces, Op. 11
Contemporary	Philip Glass — Violin Concerto no.1	Caroline Shaw — Pain & Elevation

These works are not treated as exhaustive representations of their respective periods; rather, they provide concrete examples through which the relationship between compositional structure and music-evoked thought can be examined. For example, periods characterized by high-energy, intense dynamics (e.g., Modern tracks) could elicit perceived higher acoustic energy along with shorter, less varied and less complex. Conversely, periods characterized by low-energy, reflective textures (e.g., Impressionist tracks) could elicit perceived lower acoustic energy. Rather than relying on generalized emotional valence and broad differences within the vast world of music, studies could be done to analyze fine-grained and identifiable musical variations that are easily identifiable due to existing music theory. These include, but are not limited to, the use of motif, harmonic dissonance, time signature changes, counterpoint, and robust dynamics. This study effectively continues to shift the academic focus of music psychology from whether music evokes thought to how specific musical variations affect and shape those thoughts.

V. 1. Participants

A pilot study could test these relationships using a single-group, within-subjects repeated-measures design. Approximately thirty adults between eighteen and sixty years of age would be recruited. To minimize age-related cognitive differences while maintaining diversity, recruitment will be restricted to adults aged 18 to 60 years old. Efforts will be made to ensure an even age distribution to guarantee the sample represents a broad adult demographic and is not heavily skewed towards a single age bracket. This will be achieved using a quota sampling strategy which will evenly distribute participants across three age cohorts (target $n \approx 10$ per bracket: 18–25, 26–40, and 41–60 years old), resulting in a single, well-balanced sample pool (N=30) spanning the full 18-to-60 age range. Participants would report demographic information, daily listening habits, and the extent of their instrumental, vocal, ensemble, and

music-theory training. Eligibility would be limited to people who report hearing within normal limits.

All recruitment and testing procedures will adhere to the American Psychological Association (APA) Ethical Guidelines. Participants will be required to verify their demographics, confirm they have hearing “within normal limits" (WNL), and give their informed consent to the study before participation. No prior musical training of any degree will be required; however, individual differences in musical background will be documented.

V. 2. Materials

The musical stimuli would consist of twelve instrumental excerpts: two selections from each of the six periods examined in this paper. Using two composers or works per period would reduce the risk that responses reflect an idiosyncratic reaction to a single piece rather than a broader period-level tendency. Each excerpt would last approximately two minutes, allowing enough time for the evolving structure of classical music to become perceptible without creating excessive fatigue or memory decay. Following the paradigm of van der Walle et al. (2025), all stimuli will be digitally normalized so that volume and musical variations are held constant across conditions to avoid extraneous variables.

The questionnaire will consist of a single, short, multiple-choice question (MCQ), adapted and validated by Jakubowski et al. (2024), and randomized so that participants do not see the same order every time. A prompt that asks about specific details subjects thought of while listening to the musical stimuli will also be included on this questionnaire. Some sample MCQ items are shown below:

1. Which of the following best characterizes what was going through your mind? (Select all that apply):
   [ ] A memory of a specific personal experience from my life
   [ ] A memory of media I have consumed (e.g., a movie, TV show, or video game)
   [ ] An imagined fictional story, scene, character, or scenario
   [ ] Abstract shapes, colors, or moving patterns
   [ ] Thoughts directly analyzing the music itself (e.g., the melody, harmony, tempo, etc.)
   [ ] Spontaneous thoughts about my daily life, future plans, or chores
   [ ] No internal thoughts / I was purely focusing on the sound
   [ ] Other
2. “Please describe any thoughts, mental images, or inner experiences you had while listening to this music clip.”
   [Text Box]

Additionally, to quantify results and control for variables such as familiarity and acoustic levels, an 8-point Likert scale will be administered to participants:

Measure	Scale Anchor (1)	Scale Anchor (8)
Prior Familiarity	1 = Completely unfamiliar	8 = Completely familiar
Perceived Acoustic Energy	1 = Calm / Low energy	8 = Intense / High energy

V. 2. Procedure

The main testing phase consists of twelve randomized blocks of musical excerpts to be played on psychological testing software to track responses clearly. The randomization of the prompts controls for order effects such as fatigue or practice effects. The track order will be played in different orders for each participant. Each individual trial will proceed through a rigid chronological sequence:

1. **Visual and Cognitive Fixation**: In order to eliminate external factors that could skew subjects' thoughts, the study will employ a blank gray screen with a play button in the center and no other color indications.
2. **Audio exposure**: When participants are ready to listen, they will be able to click on the play button at their own pace to start the auditory stimuli. Once started, the audio will run through the entire excerpt without pausing. Participants will be instructed to listen naturally, allowing their minds to wander and/or react freely to the audio.
3. **Thought-type response collection**: Immediately upon conclusion of the audio track, the software will present an untimed questionnaire for participants to fill out at their own pace. Participants will first respond to the short multiple-choice question and open-ended prompt.
4. **General evaluation form**: Once the qualitative questionnaire has been submitted, a short evaluation form of the tracks and listening experience will be presented using distinct 8-point Likert scales. This evaluation will collect individual ratings of personal prior familiarity (1 = never heard it before, 8 = know it perfectly) and perceived acoustic energy level (1 = calm, 8 = high energy).
5. **Stimulus and Cognitive Reset**: Once the evaluation has also been submitted, a 30-second silent rest break will follow each trial to clear cognitive lingering and auditory fatigue before moving on to the next randomized block of music.

After the completion of the twelfth final trial, the testing software will automatically transition to a debriefing statement. This statement will fully disclose the purpose of the study, revealing that the experiment specifically investigated how distinct Western classical subgenres systematically alter the cognitive pathways to either autobiographical memories or fictional narratives. The debriefing will also detail the necessity of the initial blind framing to preserve natural, unbiased cognitive responses. Participants will then be given the opportunity to ask questions regarding the study as well as reaffirm their consent to have their anonymous data included in the final analysis. They may also opt to receive a copy of the results upon the study's

completion. After receiving their gift cards and words of appreciation for their time and effort, they will be dismissed from the laboratory.

V. 3. Planned Data Analysis

To preserve methodological control and isolate the direct impact of classical subgenres on cognition, all participant data will be grouped together into a unified sample (N=30) for a single-group within-subjects, repeated measures design. The study evaluates how participants' cognitive and emotional responses shift across six distinct musical periods, which serve as the primary independent variable. The study tracks four specific dependent variables: continuous 8-point Likert ratings for prior familiarity, continuous 8-point Likert ratings for perceived acoustic energy, binary multi-response checkbox selections detailing explicit thought types (MCQ), and qualitative text responses to the prompt mapping narrative thoughtscape structures. To isolate the effects of the music itself, formal musical training will be entered into the statistical models as a continuous baseline covariate to control for participants' prior musical expertise and eliminate it as a potential confounding variable.

Inferential testing will be tailored to the distinct mathematical structures of each collected data point. For the binary multiple-choice checkboxes, a Generalized Estimating Equation (GEE) logit model will be utilized to determine if the probability of selecting a thought type changes significantly based on the musical period, while holding baseline training constant. The quantitative Likert scales for familiarity and acoustic energy will be evaluated using a one-way repeated measures. Analyses of Covariance (ANCOVA) to compare the average rating shifts across periods while controlling for training. Finally, the qualitative prompt will be examined through a mixed computational and manual analysis. To control experimental bias, two human coders blinded to the primary hypotheses will interpret the text into thematic thought types, which will be verified using Cohen's Kappa. This text will also be processed using a Natural Language Processing (NLP) framework to convert the written response into numerical text embeddings to objectively measure the vocabulary and conceptual cosine similarity in participants' spontaneous thoughtscapes across the different musical eras. A cosine similarity of 0.0 will mean that the texts share no overlapping concepts or vocabulary, and a cosine similarity of 1.0 will mean they share perfect semantic similarity. For all primary and post-hoc inferential statistical models, the criterion for statistical significance will be set at an alpha level of $p < .05$.

V. 4. Preliminary Suppositions and Implications

Three hypotheses guide this study's investigation based on contemporary frameworks of music-evoked imagery (Jakubowski et al., 2024; van der Walle et al., 2025).

First, regarding the main effect of the musical period on thought categories (H1), the structural differences presented by each period are expected to systematically drive distinct thought patterns. Specifically, the emotionally expressive textures of the Romantic era will evoke vivid fictional imagery, whereas highly structured periods (i.e., Baroque and Classical) will drive vivid autobiographical memories, and Modern tracks will more likely lead to immersive fictional

narratives. This hypothesis would be supported, for instance, by a high probability of "fictional narrative" choices during Romantic blocks, contrasted with "autobiographical memory" choices during Baroque blocks. If found to be true, this study would provide empirical evidence that classical musical periods deserve to be examined not as a monolith, but as a musical genre capable of evoking highly sensitive cognitive processing in many ways.

Second, under the familiarity-driven cognitive shifting hypothesis (H2), trials that elicit high prior familiarity scale ratings are expected to correspond to a cognitive pivot away from objective sound-tracking and toward subjective mind- wandering. Specifically, highly recognizable tracks will trigger a significantly higher probability of reporting autobiographical memories and fictional stories on the MCQ checkbox selections, along with significantly longer and more complex prompt responses, compared to unfamiliar tracks across all conditions. If H2 is supported, trials eliciting high prior familiarity ratings will correspond to a significantly higher probability of reporting autobiographical memories on the MCQ checkbox selections along with greater prompt response complexity. This outcome would demonstrate that personal listening history and subgenre exposure shift the brain away from real-time acoustic decoding and toward deeper, personal mind-wandering, like patterns seen in broader musical genres. Conversely, if H2 is not supported and unfamiliar tracks still trigger vivid personal imagery, it would suggest that the structural properties of classical music possess an inherent ability to evoke specific thought types independent of prior exposure.

Finally, it is hypothesized that the six musical periods will systematically alter both perceived acoustic energy scale ratings and response structures (H3). Specifically, periods characterized by high-energy, intense dynamics (e.g., Modern tracks) are expected to elicit significantly higher acoustic energy scale ratings along with significantly shorter, less varied and less complex prompt responses that display a higher semantic cosine similarity across the sample.

Conversely, periods characterized by low-energy, reflective textures (e.g., Impressionist tracks) are expected to elicit significantly lower acoustic energy ratings along with significantly longer, more varied and complex prompt responses that display a lower semantic cosine similarity across the sample. Before testing the primary hypotheses, a preliminary order-effects check will be conducted on the chronological trials to verify that progressive fatigue and cognitive lingering were non-significant factors ($p > .05$) due to the 30-second silent rest breaks between trials. This check will be performed using a 12-level repeated measures ANOVA.

Regarding H3, if the data analysis reveals that tracks rated higher in perceived acoustic energy yield significantly shorter and less complex prompt responses and higher semantic cosine similarity, this result will indicate that high-intensity dynamics, most prominent in Romantic and Modernist tracks, overstimulate cognitive bandwidth. Because classical tracks vary widely, it is difficult to objectively categorize entire subgenres into rigid cognitive classifications. For instance, the Romantic period contains compositions that are both quietly reflective and intensely climactic. Obtaining acoustic energy ratings trial-by-trial takes into consideration these structural nuances. On the other hand, if H3 is rejected and high acoustic energy ratings systematically

match longer, more complex prompt texts with lower semantic similarity, this would imply that high acoustic energy acts as a cognitive stimulant, enriching and diversifying the thoughtscape. Ultimately, the three hypotheses' outcomes in the proposed study will help determine whether the methodology used can successfully map how broader classical periods drive cognition and control for baseline structural and acoustic features of the musical stimuli.

V. 5. Applications of Proposed Future Research

The goal in analyzing micro-level details in music is to establish more fine-grained identifiable evidence of how the human brain perceives and navigates music to form extraordinarily vivid imagery. Hence, future research should target the countless gaps within musical genre investigation, building on top of past empirical evidence from other genres. Future research could also potentially isolate structural production shifts in the genre, such as 1980's hip-hop in relation to 2000's pop. Ultimately, by continuously exploring the micro-level data within different genres, this iterative cycle will move the framework of musical research away from treating music genres as monoliths. In doing so, the combined effort of future studies will slowly but effectively carve out a concrete and potentially universal model capable of mapping exactly how music shapes human imaginations.

The potential benefits of future studies on this topic are across both clinical and creative domains. In the realm of healthcare and therapy, such studies could contribute to the continued development of music therapy interventions. Individuals navigating through neurodegenerative conditions, such as Alzheimer's disease and various forms of dementia, often lose access to explicit memories, yet their cognitive response to music frequently remains intact (Janata). By continuing the ongoing research investigating specific compositional architectures that trigger vivid narratives and memory, this research assists in the development of personalized and meaningful treatment plans for individuals (Baird and Samson; Janata). Concurrently, within the entertainment realm, including games, films, and advertisements, future studies could offer an empirical blueprint for composers and audio engineers. Although emotional connection with the public audience ultimately stems from creative "trial-and-error" intuition, the findings of the studies can be utilized by creators to intentionally engineer more vivid and narrative-driven audio. Many of these techniques may already instinctively be applied by professionals in order to drive emotional response. However, as each individual period of Western classical music fundamentally contrasts with one another and utilizes different musical techniques, the data provided by the proposed study aims to provide a concrete, identifiable framework to aid in musical creation.

VI. Conclusion

The ongoing development of research within music psychology remains a critical investigation. By moving from the existing macro-level genre categorizations to determine cognitive outcomes to a hyperfocused genre evaluation, a more complex account of music-evoked thought emerges. The existing literature suggests that listeners' responses to music

are shaped not by genre labels alone, but by musical structure, perceived energy, and listener familiarity. Examining classical music at the level of historical subgenres therefore offers a more precise framework for understanding how music shapes thought. The same framework could be extended to subgenres and historical movements within popular, electronic, jazz, and other musical traditions, allowing researchers to move beyond broad genre categories toward the specific acoustic and cultural features that shape music-evoked thought. Therefore, future research on this topic creates promising directions for future research in both clinical and entertainment settings, such as whether particular musical characteristics are especially effective at eliciting autobiographical memories in individuals with Alzheimer's disease, dementia, or other conditions involving memory impairment, or it could offer an empirical blueprint for composers and audio engineers.

Works Cited

Audissino, Emilio. John Williams's Film Music: Jaws, Star Wars, Raiders of the Lost Ark, and the Return of the Classical Hollywood Score. University of Wisconsin Press, 2014.

Baroque Music Website. “What Is Baroque Music?” Baroque Music, www.baroque.org.

Belfi, Amy M., et al. “Music Evokes More Vivid Autobiographical Memories than Faces.” Memory, vol. 24, no. 7, 2016, pp. 975–985. https://doi.org/10.1080/09658211.2015.1061057.

Berntsen, Dorthe. “Voluntary and Involuntary Access to Autobiographical Memory.” Memory, vol. 6, no. 2, 1998, pp. 113–141. https://doi.org/10.1080/741942071.

Berntsen, Dorthe, and Nina M. Hall. “The Episodic Nature of Involuntary Autobiographical Memories.” Memory & Cognition, vol. 32, no. 5, 2004, pp. 789–803. https://doi.org/10.3758/BF03195869.

Boumpani, John, and Nicole Carteret. An Introduction to Music History and Appreciation. Open Educational Resources, 2020.

Burkholder, J. Peter, et al. A History of Western Music. 10th ed., W. W. Norton & Company, 2019.

Fink, Robert. Repeating Ourselves: American Minimal Music as Cultural Practice. University of California Press, 2005.

Grout, Donald J., and Claude V. Palisca. A History of Western Music. 9th ed., W. W. Norton & Company, 2014.

Hashim, H., et al. “Moving Pictures in the Mind: The Relationship between Visual Imagery and Narrative Construction during Music Listening.” Psychology of Music, vol. 51, no. 4, 2023, pp. 1102–1118. https://doi.org/10.1177/03057356221131456.

Hickman, Roger. Reel Music: Exploring History through Film Music. 2nd ed., W. W. Norton & Company, 2017.

Jakubowski, Kelly, and A. Ghosh. “Music-Evoked Autobiographical Memories in Everyday Life: Frequency, Selectivity, and Characteristics.” British Journal of Psychology, vol. 112, no. 3, 2021, pp. 734–756. https://doi.org/10.1111/bjop.12484.

Jakubowski, Kelly, et al. “Structural and Cultural Determinants of Spontaneous Narrative Responses to Music.” Music Perception, vol. 41, no. 3, 2024, pp. 204–221. https://doi.org/10.1525/mp.2024.41.3.204.

Janata, Petr. “The Neural Architecture of Music-Evoked Autobiographical Memories.” Cerebral Cortex, vol. 19, no. 11, 2009, pp. 2579–2594. https://doi.org/10.1093/cercor/bhp008.

Janata, Petr, et al. “Characterization of Music-Evoked Autobiographical Memories.” Memory, vol. 15, no. 8, 2007, pp. 845–860. https://doi.org/10.1080/09658210701734592.

Jarocinski, Stefan. Debussy and Impressionism. Translated by R. Myers, Macmillan, 1976.

Kluball, J., and K. Kramer. Understanding Music: Past and Present. University System of Georgia, Fine Arts Open Textbooks, 2019.

Margulis, Elizabeth H., et al. “Narratives Constrained by Music: A Cross-Cultural Study.” Proceedings of the National Academy of Sciences, vol. 119, no. 28, 2022, e2201580119. https://doi.org/10.1073/pnas.2201580119.

Metzer, David. Musical Postmodernism. Cambridge University Press, 2009.

Morgan, Robert P. Twentieth-Century Music: A History of Musical Style in Modern Europe and America. W. W. Norton & Company, 1991.

Palmer, Christopher. Impressionism in Music. Charles Scribner's Sons, 1973.

Pasler, Jann. “Impressionism.” Grove Music Online, 2001. https://doi.org/10.1093/gmo/9781561592630.article.13740.

Ross, Alex. The Rest Is Noise: Listening to the Twentieth Century. Farrar, Straus and Giroux, 2007.

Schoenberg, Arnold. Style and Idea: Selected Writings of Arnold Schoenberg. Edited by L. Stein, University of California Press, 1978.

Smallwood, Jonathan, and Jonathan W. Schooler. “The Science of Mind-Wandering: Empirically Navigating the Stream of Consciousness.” Annual Review of Psychology, vol. 66, no. 1, 2015, pp. 487–518. https://doi.org/10.1146/annurev-psych-010814-015331.

Taruffi, Liila, et al. “Effects of Sad and Happy Music on Mind-Wandering and the Default Mode Network.” Scientific Reports, vol. 7, no. 1, 2017, article 14396. https://doi.org/10.1038/s41598-017-14849-0.

Taruskin, Richard. Music in the Nineteenth Century: The Oxford History of Western Music. Oxford University Press, 2010.

Tymoczko, Dmitri. A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice. Oxford University Press, 2011. https://doi.org/10.1093/acprof:oso/9780195336672.001.0001.

van der Walle, H. A., et al. “Thoughtscapes in Music: An Examination of Thought Types Occurring during Music Listening across 17 Genres.” Psychology of Music, 2025. https://doi.org/10.1177/03057356251346654.

How do Social Media Platforms Transform Fandom Rituals Into Modern Forms of Collective Practice? By Siyan Wang

Acknowledgements

Thank you for the guidance of mentor Vinzent Wesselmann from the University of Oxford in the development of this research paper.

Abstract

This study investigates how social media platforms restructure ritualistic practices within Chinese teenage K-pop fandoms. Using a mixed-methods design combining a quantitative survey (N=70) and semi-structured interviews (N=5), the research examines shifts from traditional, media-dependent fandom to highly interactive digital fan practices. Findings show that platform features such as hashtags, ranking mechanisms, and real-time notifications intensify participation frequency and transform casual fandom into structured collective actions. The study concludes that social media does not merely mediate fandom activities but actively restructures fan rituals by imposing new metrics of visibility and behavioral expectations.

INTRODUCTION

Over the past decade, the way young Chinese K-pop fans interact with their idols has undergone significant changes. K-pop idols, who are highly trained artists in the Korean music industry, typically debut in groups or as soloists under major agencies that manage their training, image, and careers. World-famous groups like BTS and BLACKPINK are known for catchy music, synchronized choreography, and strong fan engagement.

K-pop was once accessed by Chinese fans mainly through traditional mass media. Before the rise of social media, most high school fans learned about idols through TV music programs, entertainment magazines, imported CDs, and limited information shared by early fan clubs. These interactions were largely one-directional and passive. Earlier studies comparing pre-social media and social media environments suggest that before the rise of interactive platforms, fan engagement in China was limited by the communication technologies available at the time (Jia et al., 2021; Yin and Xie, 2018). As social media platforms emerged, these limitations were overcome, enabling fans to connect, organize, and interact more frequently and directly than in the traditional media era.

Since the early 2010s, social media has become central to youth online culture in China. By 2023, China had 1.079 billion internet users, including 191 million adolescent users (CNNIC, 2023). Chinese teenage K-pop fans began to shift their fandom activities from offline events and top-down mass media consumption to highly interactive and democratic online spaces. Daily engagement, communication, and collective activities were constructed through platform features and algorithms rather than physical gatherings and offline engagement (Yin, 2020).

Weibo became one of the most influential platforms for fan activity in China during the early 2010s. By December of 2025, Sina Weibo had already accumulated 567 million monthly active users and 252 million daily active users (Weibo Corporation, 2025). Its hashtags,

algorithm-based content distribution, real-time notifications, and interactive functions enabled fans to communicate instantly and continuously. These features transformed fan participation from occasional, offline, and media-dependent consumption into frequent, structured, and highly visible online behaviors.

Since this shift, Chinese K-pop fandom has developed highly organized practices. Many fan communities create data support teams to coordinate daily tasks like increasing music-chart streams, boosting social media engagement, and coordinating collective posting schedules in order to increase an idol's visibility and popularity. "Super-topic" (超话 chaohua) further gamifies participation by rewarding sustained engagement with visibility and status (Chen, 2024). These design features create an environment where young fans perform similar tasks daily, forming shared digital routines.

Despite the highly structured and labor-intensive nature of these activities, fans find profound meaning and emotional refuge in these digital spaces. A 17-year-old participant described her daily routine as a strict regimen of "checking into Super-topics, reposting promotional content, and boosting music video views" before bed. Yet, she emphasized that this labor does not feel burdensome. Instead, interacting with like-minded peers provided a sense of belonging and a relaxing escape from academic pressure. This paradox highlights how platform-enabled rituals, despite their rigidity, function as a vital social space where young fans can alleviate stress and find solace.

Previous studies have explored several related topics, including how K-pop music spreads globally through digital networks, how fans express their identities and sense of belonging on the internet, and how continuous media exposure strengthens parasocial relationships (Jung, 2014; Gerbner and Gross, 1976). However, less attention has been paid to how platforms like Weibo reshape the structure of fan behaviors, especially among Chinese K-pop fandom groups. Existing academic research often focuses on the psychological attachment of global fan groups or individuals rather than the collective and regular behavioral patterns that define fan rituals.

In China, these collective rituals are deeply rooted in a socio-cultural soil that emphasizes collectivism, group cohesion, and disciplined conformity, which are also shaped by institutional environments like Chinese secondary schooling (Fung, 2019; Zhang & Negus, 2020). Combined with Weibo's visibility metrics, ranking systems, and coordination tools, this cultural background enables fan communities to organize highly disciplined forms of "data labor" or "data fans" (Yin, 2020; Zhang & Negus, 2020). This study therefore explores the following question: **How do social media platforms like Weibo transform traditional K-pop music fan rituals into modern collective practices for Chinese high school students?** Through survey data and interviews, this study analyzes how students participate in daily fan activities, how platform features shape repetitive collective behaviors, and how online participation compares to previous offline rituals.

LITERATURE REVIEW

Existing research on fandoms, media participation, and celebrity culture provides a solid foundation for understanding how fans interact with idols and with each other. Scholars have examined fandom from multiple dimensions: rituals and collective practices, the relationships fans develop with idols and each other, and the influence of digital platforms on fan behavior and fandom compositions. Although these studies provide important insights, they often focus upon Western contexts and broader fan communities, rather than the specific experiences of teenage K-pop fans in China. In addition, many papers discuss fan behavior as a broader concept but do not focus on how platform-specific features—such as those of Weibo, a Chinese social media platform—shape the everyday routines and behaviors of young fans. The following literature review surveys the existing scholarship across five key areas and identifies the gaps that this study aims to fill: fandom rituals, collective practices, fan-idol relationships, and media participation.

Fandom Rituals

Durkheim defines rituals as repeated and structured actions that help a group of people feel connected through collective emotional links (Durkheim, 1912). This sociological idea can be applied to modern fan groups, whose activities are often repetitive, collective, and structured. For example, many K-pop fan behaviors—such as celebrating an idol's birthday, posting in a certain format, or participating in organized activities—can be seen as ritual-like practices.

Media research also helps explain why repeated fan engagement may become ritualized. Gerbner and Gross (1976) state that television can cultivate audiences through repetitive patterns, shaping their conceptions of reality, which thus affects their everyday behavior. Moreover, emotional investment in celebrities can develop into parasocial relationships, in which fans experience a sense of intimacy with public figures through ongoing media contact (Redmond, 2018; Xu et al., 2023). These studies suggest that when fans repeatedly consume media about their idols or related content, they may develop consistent online routines or emotional responses, which can lead to ritual behaviors. Empirical studies on K-pop fandoms further demonstrate how routine online activities—such as daily news checking, streaming music videos, and community chatting—can reinforce group identity through recurring patterns of participation (Jung, 2014).

However, existing research paid limited attention to how Chinese social media platforms structure these rituals. For example, Weibo introduces highly structured daily tasks, such as checking in on Super-topics, re-posting certain posts, or participating in certain comment formats, that shape how fans participate. On Weibo, practices such as checking in to Super Topic communities, reposting coordinated content, and using shared hashtags or comment formats can turn participation into visible and measurable daily tasks. These practices combine repetition, symbolic communication, and collective goals, but the role of platform-specific mechanisms in organizing them remains underexplored. This study addresses this gap by examining how Weibo shapes ritual-like practices among Chinese high-school K-pop fans.

Collective Practice

Durkheim's idea of collective life suggests that groups create shared emotional experiences through repeated actions, maintaining a sense of unity (Durkheim, 2008). In applying this to digital culture, Mark Deuze argues that people today live in a "media life," where media are environments we live in rather than merely tools we use (Deuze, 2011). In online fandoms, fan behavior is rarely isolated; instead, fans participate in collective practices as both "prosumers"—producers and consumers—creating shared cultural environments and identities.

In this study, collective practices are specifically defined as algorithmically organized, highly synchronized online fan behaviors, such as coordinated reposting, collective comment control (*kongping*), and systematic voting/ranking (*dabang*). Global K-pop studies show that such participation relies heavily on digital platforms to maintain community structures (Jung, 2014). On Chinese social media like Weibo, this mechanism is further intensified by features such as Super-topic (*chaohua*) rankings. These rankings can turn daily fan participation into competitive collective activity, as fans may feel encouraged to maintain their idols' visibility and popularity.

However, while existing literature describes how media environments support collective participation, it relies heavily on earlier platforms like Twitter or Facebook. Scholars rarely address how platform convergence in China—where ranking, community, publicity, and metrics merge into a single interface—produces stronger forms of collective pressure and coordinated fan labor. Thus, the gap lies in analyzing collective practices within these advanced, metric-driven social media ecosystems.

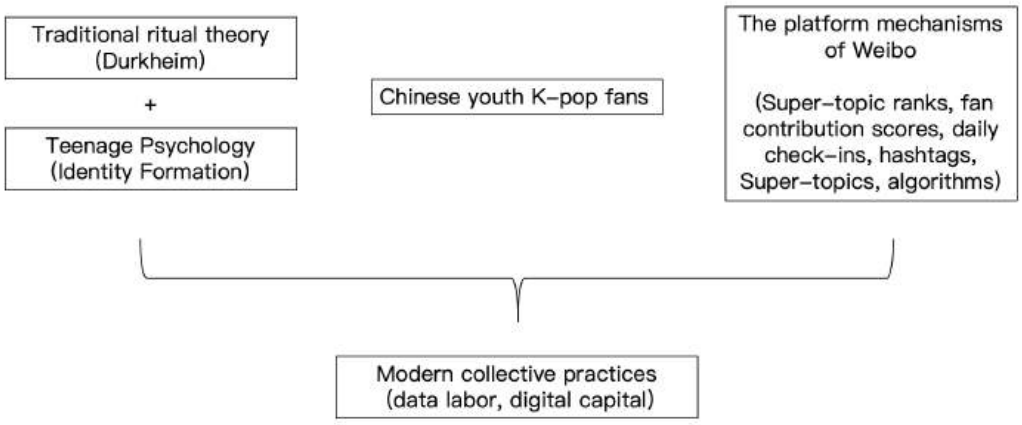


Figure 1: mind map for the formation of modern collective practices

Fan-idol relationship

Consistent media contact can foster habitual emotional responses to media figures (Gerbner and Gross, 1976). More recent research looks directly at how people form emotional bonds with celebrities online. Xu et al. find that digital platforms allow fans to feel more personally connected with their idols, even without direct communication and interaction (Xu et al., 2023). This helps explain why K-pop fans often have strong feelings of engagement. Platforms make idols appear closer, more accessible, and more real in everyday life.

Jung (2014) points out that the image of K-pop music idols is presented through stories of hard work, continuous progress, and emotional engagement, which makes them highly attractive. This has made fans view their support for idols as a meaningful and morally valuable act. Therefore, fans often express their sense of responsibility for the idol's success. This finding explains why many fan behaviors become sustained, ritualistic, and collective.

However, there is limited research on how platform functions specifically shape the intensity and continuity of parasocial ties for younger users. Unlike adults who possess more stable self-identities and emotional regulation, high school students are in a critical stage of identity formation, rendering them more vulnerable to platform algorithms that incentivize deep emotional dependency. Furthermore, existing studies rarely address the dark side of these parasocial ties among adolescents, particularly the emotional distress and parasocial dissolution experienced when idol image collapses occur (Hu, 2023). This creates a gap in understanding how platform features shape fan-idol connections among teenagers in China.

Media Activities

Deuze's concept of "media life" is helpful for understanding fandom because young fans' activities, whether browsing Weibo, watching livestreams, or joining group chats, are embedded in their everyday routines rather than occurring as separate or occasional behaviors (Deuze, 2011). Similarly, Jan-Erik Nordlund's work suggests that online activity is not only shaped by social intentions but also by the design of the platform itself, which guides how users engage with content and with each other (Nordlund, 1978). This perspective is especially relevant for studying K-pop fandom on Chinese platforms like Weibo, where algorithms, visibility mechanisms, and interactive features shape both what fans see and how they respond.

Building upon these frameworks, recent research highlights how algorithmic engagement and the data-ization of fandom drive fans to voluntarily become "digital laborers" (Yin, 2020). Yin explains that fans willingly perform uncompensated data labor because platforms transform fan affect into quantifiable metrics. In this context, task-based participation, such as mass reposting and comment control, is framed as a shared emotional responsibility to increase visibility for their idol while generating collective achievement and group solidarity. On Weibo, engagement is constantly quantified and displayed through public metrics such as Super-topic ranks, fan contribution scores, and task completion points. By gamifying participation and datafying fan affect, the platform aligns fans' self-worth with visible metric outputs, transforming routine media consumption into self-motivated, data-driven digital labor (Yin,

2020). Based on these insights, this study hypothesizes that Weibo's mechanics actively structure high school fans' daily routines into gamified, quantified rituals.

Summary

Although existing studies provide useful frameworks for understanding fandom, celebrity culture, and digital participation, several research gaps remain. First, there is limited research on K-pop fandom in China, especially among teenage female fans. Second, most studies discuss fan rituals in general terms, but few examine how platform-specific features actively structure everyday rituals. Third, existing literature rarely connects online fan activity with offline practices such as birthday café events. This study addresses these gaps by examining how Weibo's platform features shape the everyday online and offline fan practices of Chinese teenage K-pop fans.

METHODOLOGY

Research Design

The methodology combines quantitative survey data with a qualitative analysis of semi-structured interviews. Using both methods allows the study to capture broad behavioral patterns while also gaining deeper insights into how individual fans interpret their own practices and how social media structures their daily fandom routines. The survey identifies general trends in media use, fan activities, and participation frequency, whereas the interviews offer more detailed explanations of motivation, ritual meaning, and platform influence.

Participants

Of the 70 survey respondents, 53 identified as K-pop fans. Among these fans, 36 (67.9%) are high school students aged 16-18 years old. Within this age group, 31 participants identified as female, four as male, and one preferred not to disclose their gender. This group was selected because high-school students are active participants in contemporary youth fandom culture and frequent users of platforms such as Weibo and Douyin. Five high school students were chosen for follow-up interviews based on their survey responses regarding daily social media routines and fan engagement levels.

All interview participants were recruited from the same high school through classroom announcements and school networks. Survey respondents, however, were recruited through personal WeChat networks and K-pop-related Weibo communities. This approach ensured that the sample included both classmates within the same institutional setting and broader fan participants from social media. However, as a non-probability convenience sample that was predominantly female, the findings may overrepresent highly active fan circles and may not be generalizable to all Chinese high school K-pop fans.

Data Collection Methods

Survey

The survey questions aimed to evaluate fans' media use habits, their attitude towards online and offline participation, their thoughts on media's role in fan activities, and their emotional connections with fans, idols, and influencers. The survey was designed using Wenjuanxing, distributed online through personal networks on WeChat and Weibo, and took approximately 3-5 minutes to complete.

Interview

Following the survey, semi-structured interviews were conducted to explore participants' experiences and understandings of online and offline fan rituals. The interview questions were organized around four main themes: participants' experiences of entering K-pop fandom, their social media use and daily fandom activities, repeated and collective practices within their fandom communities, and the relationship between online and offline participation and the meaning of "being a fan." The questions were designed to explore how participants understood their own fandom practices and how these practices were influenced by social media platforms. Each interview lasted around 40 minutes and was conducted in person. With participant consent, interviews were recorded and transcribed for further analysis.

The five interview participants were selected from survey respondents who indicated that they were willing to participate in a follow-up interview. The selection aimed to include participants with different levels of fandom participation, different patterns of platform use, and different grade levels, allowing the interviews to capture a range of experiences within the sample.

Procedure

The study was conducted in two stages. First, the survey QR code was distributed to participants and responses were collected over a period of three days. After reviewing survey data, participants who indicated that they were willing to be interviewed were contacted individually. Interviews were scheduled and conducted using the same set of guiding questions to ensure consistency across sessions, although follow-up questions varied according to each participant's responses. Transcriptions were anonymized before analysis.

Ethical Considerations

The study's protocol was approved by the school supervisor. Prior to data collection, written student assent and parental consent were obtained. Participation in all surveys and interviews was strictly voluntary, with participants informed that they could withdraw at any time without consequence. To ensure confidentiality, all personal identifiers were anonymized. Interview audio recordings were stored securely on a private, password-protected mobile device with no external sharing, and all raw media files will be permanently deleted three months after the completion of this paper.

Data Analysis

Quantitative Analysis: Survey data were analyzed using descriptive statistics (frequencies and percentages) in Wenjuanxing to calculate response patterns, such as the frequency of daily Super-topic check-ins, time spent on Weibo, and participation in fan tasks.

Qualitative Analysis: Interview transcripts were analyzed using Braun and Clarke's (2023) thematic analysis. The analytical process followed five iterative steps: (1) familiarizing with the transcript data; (2) generating initial codes (e.g., "algorithmic mechanics," "daily check-in routine," and "parasocial attachment"); (3) searching for broader themes; (4) reviewing and refining themes against the raw data; (5) defining and naming key themes regarding digital fan rituals.

Integration: This study used an explanatory sequential mixed-methods design. Quantitative findings identified broad patterns of participation patterns, while qualitative interview insights were integrated to explain the underlying emotional motivations and daily social media routines. The two forms of data were integrated in the Results and Discussion sections to examine how Weibo's platform features shaped repeated collective fan practices.

RESULTS

This section presents the findings from the survey of 70 Chinese teenage K-pop fans, among whom 53 people self-define as K-pop fans. The results are divided into four categories: platform usage and participation frequency, media participation, community interaction, and views on ritualistic behaviors.

1. Platform Usage and Participation Frequency

Among all the respondents, Weibo was the platform that held an absolute dominant position in their social media usage. 86.79% of respondents (n=53 K-pop fans) chose Weibo as the main channel for obtaining K-pop information, participating in fan activities, or following the updates of their idols. This was followed by TikTok (56.6%) and Instagram (33.6%).

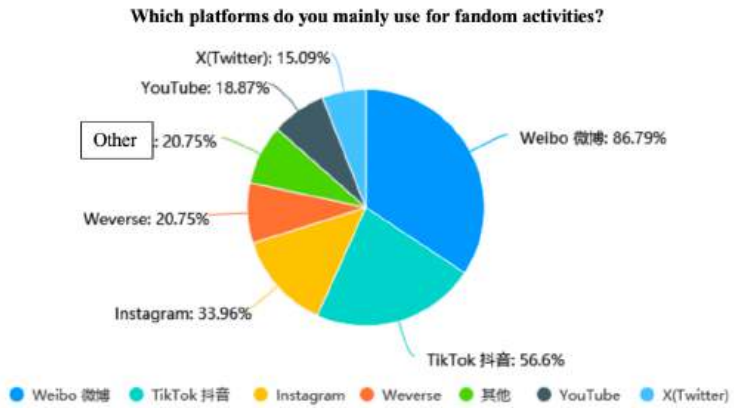


Figure 1: Responses to survey question "Which platforms do you mainly use for fandom activities?" (multiple answers possible)

This distribution pattern is associated with the overall popularity of Weibo among Chinese users, its large daily active user base, and its specially designed features such as Super-topics, real-time hot list rankings, and structured "data support" functions. The unique design of these platforms is likely to have solidified Weibo's position as the primary platform for Chinese K-pop fans to participate in K-pop fandom activities.

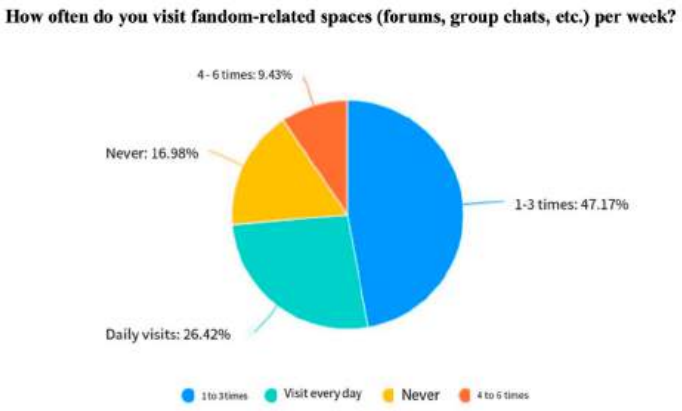


Figure 2: Responses to survey question “How often do you visit fandom-related spaces (forums, group chats, etc.) per week?”

In terms of engagement frequency, 47.17% participants participate in fandom activities 1–3 times per week, while 26.42% engage daily.

The prevalence of the frequency of "1 to 3 times per week" may reflect the rhythm of teenagers' daily lives, such as school schedules, homework loads, and limited free time. These factors may limit sustained participation, but these activities still have a high degree of structure, periodicity, and repetition, and thus can be regarded as a ritualized behavior. Although this survey cannot provide an explanation for the causal relationship, a reasonable speculation is that fans will schedule their fan activities during predictable leisure periods such as commuting time, lunch break time, or evening time. Future research may be able to further explore the relationship between daily activities and the time spent on online K-pop fandom activities.

2. Media Participation

Contrary to the initial projection that most participants only browse the idol-related contents, the data indicates that the majority of participants actively “like or comment”(81.13%), while 67.92% of the participants only browse without producing any content.

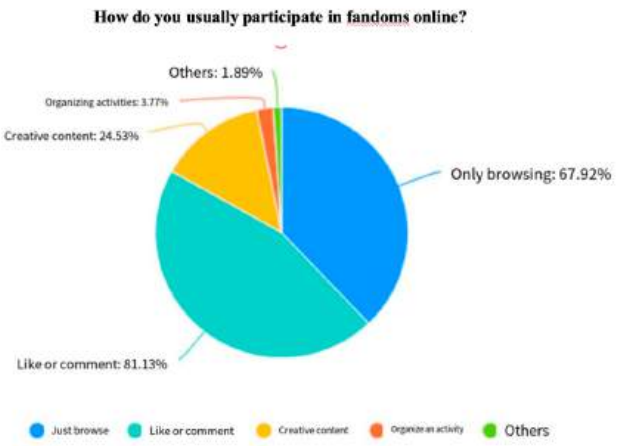


Figure 3: Responses to survey question "How do you usually participate in fandoms online?" (multiple answers possible)

This suggests that fans commonly engage in light, low-effort participation, but they are not entirely passive. The relative ease of "liking" or leaving short comments, especially in fandom contexts where collective commenting (控评, kongping), boosting posts, or dropping specific hashtags are common. This may explain why users participate in these behaviors even though browsing requires less effort. Furthermore, this low-effort participation can "be rewarded with prestige and reciprocity in the fan community", as they help in "artistic production, gate-keeping, curating, cataloging, editing, scheduling, distributing, marketing/advertising, and retailing" (Hou, 2025; Ming and Leung, 2017).

Some participants described collective commenting (控评, kongping) as a shared fan practice involving coordinated comments, reposts, fixed phrases, and emojis. For instance, fans use fixed phrases such as "the no. 1 dancer of the 5th generation" (五舞一, wuwuyi), unified emojis, mass commenting and re-posting those comments to form a "comment wall." These comment walls are useful because the algorithm recognizes them and distributes the content to broader audiences. Collective commenting not only serves informational functions but also symbolic ones, reinforcing the collective consciousness within fandom groups. This specific micro-ritual may lower the psychological barrier to participation and encourage users to engage in visible, supportive acts rather than remain entirely silent.

In contrast, participation patterns in longer-duration activities, such as creating original fan art, producing edits, or organizing events, were lower. This data reveals participation inequality in digital fan communities, broadly consistent with Nielsen's (2006) "90-9-1 rule" of online engagement. While the vast majority of respondents (e.g., 81.13%) engage in low-effort, highly accessible daily rituals like liking or checking in, only a small active minority contributes to time-intensive content creation and event organization.

3. Community Interaction

The survey results reveal differences in how participants formed relationships through fandom activities. For online interactions, 50.94% of respondents reported not making friends through online fandom spaces, while 49.06% indicated that they had formed online friendships. This suggests a near-even split in reported online friendship formation.

In contrast, 60.38% of participants reported not making friends through offline meetups or events, compared with 39.62% who had. Therefore, more respondents reported forming friendships through online fandom spaces than through offline activities.

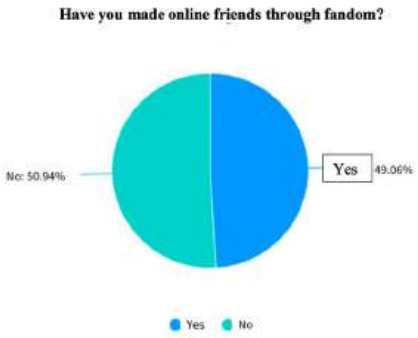


Figure 4: Responses to survey question “Have you made online friends through fandom?”

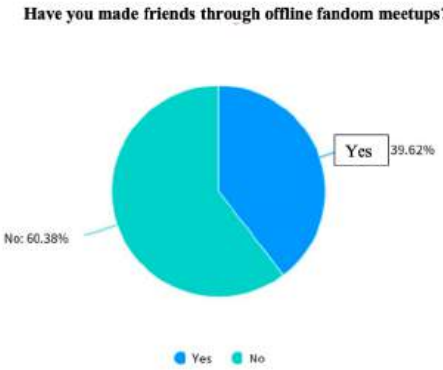


Figure 5: Responses to survey question “Have you made friends through offline fandom meetups?”

Online and offline participation also differed in their practical conditions. Online fandom activities, such as commenting, liking content, and participating in Super-topics, can take place without physical co-presence. In contrast, offline fandom activities require participants to be present at the same time and place, which creates a higher participation barrier for the respondents in this study.

Overall, the survey indicates that online friendship formation was slightly more common than offline friendship formation among the participants, while both forms of social connection were present within the fandom.

4. Views on Ritualistic Behaviors

Participants were asked to identify which fandom activities they considered "rituals," and the responses reveal clear patterns in how teenagers interpret symbolic and repeated practices within K-pop communities.

Figure 6: Responses to survey question "Which of the following do you see as 'rituals' in your fandom life?"
(multiple answers possible)

A significant majority of participants (77.36%) indicated that watching every release of music videos and albums immediately can be considered as a ritual. This perception can reflect a shared expectation of timely participation, in which fans treat the immediate viewing of music videos, teasers, or live performances as a collective routine that expresses loyalty and maintains consistency with the norms of the fandom group.

75.47% of participants classified participation in birthday or anniversary projects as ritualistic. These projects typically take place annually, involve well-organized preparation, and carry strong symbolic meaning within fandom culture. The label of "ritual" suggests that participants recognize these even if not simply as celebratory activities but as structured, repeated practices that celebrate a significant moment for the idol.

54.72% of participants also viewed meeting friends to celebrate idol events as a ritualistic behavior. Unlike the aforementioned platform-based activities, the offline gatherings require face-to-face mutual participation, reflecting the social nature of the fan group. The participants seemed to consider these activities as having ritual or commemorative qualities, even when carried out in an informal setting.

Overall, the findings suggest that teenagers recognize both platform-mediated activities and offline, socially coordinated activities as rituals. What makes these actions "ritualistic" is not merely their frequency but their symbolic weight, shared expectations, and role in structuring the fandom life of everyday fan participation.

ANALYSIS

The research question of this study is, **"How do social media platforms transform fandom rituals into modern forms of collective practice?"** The survey results and interview responses demonstrate that digital platforms, especially Weibo, help reshape ritual participation by encouraging repetitive behaviors, enabling collective actions, and reorganizing the relationship between online and offline forms of fan engagement. This section analyzes how these mechanisms play a role in modern teenage K-pop fandom among Chinese high school female students by connecting the results with existing theories and the interview responses I have collected.

1. Repetitive Digital Behaviors as Platform-Structured Rituals

Durkheim (2008) conceptualizes ritual as repeated, symbolically charged actions that reinforce group belonging and emotional solidarity. In digital contexts, Xu et al. (2023) argue that online rituals emerge when platforms "stabilize repeated participation through interface design, symbolic incentives, and algorithmic reinforcement." The findings are consistent with the idea that repeated participation can take ritual-like forms.

Rather than simply reflecting personal preference, the high frequency of liking, commenting, and checking Weibo suggests that these behaviors are continuously reinforced by the platform itself. Weibo's notification system, algorithmic recommendations, and Super-topic reminders repeatedly direct users back to idol-related content, embedding fandom participation into everyday routines. This mechanism is reflected in both the survey and interview findings. The data show that 81.13% of participants regularly like or comment on posts, and 47.17% engage in fandom activities 1–3 times per week, while 26.42% do so daily. Interviewees repeatedly described these activities as habitual. Interviewee 1 mentioned, "I check Weibo every morning before school when I brush my teeth; it's just part of my routine." Interviewee 3 similarly explained, "Whenever I feel bored during breaks, I naturally open Weibo to see if anything new has happened to my idol."

These findings support Durkheim's concept of ritual as repeated practices that reinforce collective belonging (Durkheim, 2008). However, unlike traditional rituals, the repetition observed here is not generated solely through participants' intentions. Instead, it is stabilized through platform mechanisms. By continuously reminding users to return, Weibo shapes ordinary media use into habitual ritual performances that become embedded in daily life. Therefore, the platform appears to play an organizing role in participants' routine fan activities.

2. Platform Features Enabling Ritual Behaviors

The results also highlight Weibo's status as the dominant platform for fandom action: 86.79% of respondents named it their primary site of participation. This dominance is not just a reflection of popularity but also demonstrates how Weibo's unique platform features help enable collective, synchronized rituals.

(a) Super-topic Sign-ins

Rather than functioning as a simple community platform, the Super-topic system acts as a mechanism that shapes user participation by encouraging users to return every day to achieve daily check-in records in order to maintain community rankings. Interviewee 4 described this as a routine normative practice: "If I don't sign in, the level of our Super Topic drops. Everyone checks in every day—it's just something we all do." Therefore, daily sign-ins can standardize ritual frequency and create a measurable, collective routine. This shapes participation into a habitual obligation rather than a voluntary activity, illustrating how the platform regulates ritual frequency. This process also reflects the domestication of technology, whereby digital technologies gradually become embedded in users' everyday routines and are experienced as ordinary rather than exceptional (Silverstone et al., 1996). Instead of consciously deciding whether to participate each day, fans gradually internalize the platform's check-in system as part of their daily habits.

(b) Collective Commenting and Standardized Comment Formats

Collective commenting demonstrates how platform mechanisms shape fan expression rather than simply enabling communication. Fans make efforts to control the comment section through collective commenting, which takes form through the mass-posting of supportive messages, specific emojis, or fixed phrases. Interviewee 1 explained, "We comment in the same format so it looks neat and boosts our idol's image." Interviewee 5 mentions that "we believe this can increase the exposure of our idol to the public." This creates a structured vocabulary used in K-pop fandom ritual practices. Alongside expressing personal opinions, fans may reproduce standardized symbolic language that strengthens collective identity, consistent with Durkheim's understanding of ritual as the production of shared meanings. At the same time, these repetitive commenting practices resemble forms of digital labor, as users continuously generate engagement and visibility that benefit both the platform and the idol's online presence (Terranova, 2012).

(c) Hot Search Rankings

Hot Search Rankings illustrate how platform algorithms shape users' perceptions of visibility and success. Interviews showed that participants perceived that their actions directly affected the idol's visibility. As interviewee 5 noted, "If we don't keep posting in the tag, our idol disappears from the page." This perception encourages fans to engage in repetitive collective actions, such as mass re-posting at strategic times. For example, during new songs' release, fans will post massive content under their idols' tag in order to help the tag appear in Hot Search Rankings. The act of re-posting becomes ritualistic because it is performed collectively, repetitively, and with a shared belief in symbolic effect. Also, because visibility is treated as something fans can collectively influence, algorithmic rankings shape repeated reposting into a meaningful ritual rather than a purely technical action.

(d) Quantification of Participation

Quantification represents perhaps the clearest example of how the platform makes participation visible, comparable, and potentially competitive. Indicators such as the number of reposts, the record of continuous check-in days, and participation rankings make this ritual behavior clear, measurable, and competitive. This shapes the ritual from a symbolic act into a quantifiable and visible contribution, thereby enhancing the sense of responsibility and achievement among fans. Together, these features demonstrate that Weibo does not merely host fan rituals, but also shapes and enhances them, turning individual media habits into coordinated collective practices that reinforce fandom group identity. Participation, therefore, becomes a form of digital labor, where fans continuously contribute measurable engagement in exchange for symbolic recognition and the perceived success of their idols.

3. Online vs. Offline Rituals: The Platform as the Primary Ritual Space

The results also show how platforms reorganize the spatial dimension of fandom rituals. Survey data indicates that online friendships are almost evenly split, while offline friendships are less common in the sample. This suggests that online interaction is easier and more accessible, making digital platforms the main ritual space.

Interview responses further explain this distinction. Interviewee 2 described a birthday café event as an example of offline gatherings: "The atmosphere offline is so engaging—you scream together and feel the same excitement. It makes everything more real." Interviewee 3 mentioned a concert live-stream viewing event: "Being in the same room with fans who understand you feels powerful, but I can't do it often because of school." These quotes show that offline gatherings generate a stronger sense of communitas because participants share physical space, emotional synchrony, and embodied experiences. This aligns with the notion of "communitas" (Turner, 2004) and Collins' theory of interaction ritual chains, which argues that physical co-presence and shared emotional attention generate stronger emotional energy and group solidarity than mediated interactions (Collins, 2004).

However, although online rituals (such as daily check-ins, synchronized reposting, and collective commenting) may be less emotionally intense, they are more sustainable because platform infrastructures remove barriers of time and location. As a result, ritual intensity is partially replaced by ritual frequency. Interviewee 4 said that "I can join online activities anytime, even during a 10-minute break."

Thus, social media platforms shift the main platform of ritual life, making digital rituals frequent, structured, and collective because it is a more convenient alternative choice for fandom activities compared with occasional, high-impact offline rituals. Rather than replacing traditional rituals, social media redistributes ritual participation from occasional, high-intensity physical gatherings to frequent, platform-mediated collective practices.

LIMITATIONS

This study has several limitations. First, the survey used a small non-probability convenience sample, while the five interview participants were recruited from the same high

school. The sample was also predominantly female and may overrepresent highly active K-pop fans. Therefore, the responses collected may not be generalized to all Chinese teenage K-pop fans, particularly male, gender-diverse, less active, or geographically diverse fans.

Furthermore, this study mainly focuses on Weibo and does not examine cross-platform practices on Bilibili, Douyin, WeChat, Weverse, or other international platforms. Therefore, the conclusions are most applicable to online environments similar to Weibo. Moreover, the reliance on self-reported data might have potential biases, including selective memory and social expectation responses. Broad terms such as "online participation" may have different interpretations by the participants. Overall, these findings should be regarded as exploratory. Future research should adopt larger sample sizes, more diverse samples, clearer operational definitions, and cross-platform comparisons to produce more robust findings.

CONCLUSION

This study examines how Chinese high school students participate in K-pop fandom groups and analyzes how platform design may help organize repeated and coordinated forms of fan participation. Common digital behaviors such as liking, sharing, commenting, viewing the rankings of hot topics, and watching live-streams may gradually acquire ritualistic characteristics due to their repetitiveness, coordination, and shared symbolic meanings. Interviews also highlighted the role of check-ins, notifications, and visible participation metrics.

Participants believed that coordinated engagement could increase their idols' visibility, thereby reinforcing a sense of collective responsibility. These trends suggest that platform features may help coordinate individual actions into shared fan practices.

While daily fandom is dominated by digital engagement, offline events continue to hold emotional significance for some participants. Rather than replacing traditional practices, social media appears to reshape fandom into a hybrid ritual model—a blended ecosystem where daily online repetition and periodic offline emotional intensity may reinforce one another. Beyond the specific case of K-pop fandom, these findings may contribute to broader discussions of digital ritualization.

The study illustrates how technology can regulate social behavior: algorithmic incentives and quantification systems may shape casual scrolling into structural, meaningful collective action. Future research could use longitudinal and cross-platform designs to examine how digital fan rituals develop over time and how intensive fan participation relates to young people's well-being and offline relationships. In conclusion, this study suggests that Chinese K-pop fan rituals are shaped by both platform architecture and collective fan agency. This hybrid ecosystem highlights a continuous interplay between algorithmic spaces and physical environments, demonstrating how social media restructures, rather than simply facilitates, modern collective practices.

Works Cited

Braun, Virginia, and Victoria Clarke. Thematic Analysis. Washington, DC: American Psychological Association, 2023.

Chen, Tiancheng. Understanding Trans-semioticising Expressions on Chinese Social Media. PhD diss., University College London, 2024. Accessed September 2, 2026. https://discovery.ucl.ac.uk/id/eprint/10186241.

Collins, Randall. Interaction Ritual Chains. Princeton, NJ: Princeton University Press, 2004. Accessed September 2, 2026. https://doi.org/10.1515/9781400851744.

Deuze, Mark. "Media Life." Media, Culture & Society 33, no. 1 (2011): 137-48. Accessed September 2, 2026. https://doi.org/10.1177/0163443710386518.

Durkheim, Emile. The Elementary Forms of Religious Life. Mineola, NY: Dover Publications, 2008. Originally published 1912.

Fung, Anthony Y. H. "Fandom, Youth and Consumption in China." European Journal of Cultural Studies 12, no. 3 (2009): 285-303. Accessed September 2, 2026. https://doi.org/10.1177/1367549409105365.

Gerbner, George, and Larry Gross. "Living with Television: The Violence Profile." Journal of Communication 26, no. 2 (1976): 172-99. Accessed September 2, 2026. https://doi.org/10.1111/j.1460-2466.1976.tb01397.x.

Hou, M. Language, Celebrity Fandom, and Political Activism in Chinese Media: Battling for Idols Online. 1st ed. London: Routledge, 2025. Accessed September 2, 2026. https://doi.org/10.4324/9781003275589.

Hu, Mu. "Parasocial Relationship Dissolution and Deterioration." In The Oxford Handbook of Parasocial Experiences, edited by R. T. Forster, 147-70. Oxford: Oxford University Press, 2023. Accessed September 2, 2026. https://doi.org/10.1093/oxfordhb/9780197650677.013.6.

Jia, B., J. Li, and J. Ma. "Transformation of Fan Culture under the Influence of Social Media." In Proceedings of the 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021), 2173-78. Paris: Atlantis Press, 2021. Accessed September 2, 2026. https://doi.org/10.2991/assehr.k.211220.375.

Jung, Sun. Korean Masculinities and Transcultural Consumption. Hong Kong: Hong Kong University Press, 2010. Accessed September 2, 2026. https://doi.org/10.5790/hongkong/9789888028672.001.0001.

Jung, Sun. "Youth, Social Media and Transnational Cultural Distribution: The Case of Online K-pop Circulation." In Mediated Youth Cultures: The Internet, Belonging and New Cultural Configurations, edited by Andy Bennett and Brady Robards, 114-29. Basingstoke: Palgrave Macmillan, 2014. Accessed September 2, 2026. https://doi.org/10.1057/9781137287021_8.

Ming, Y., and L. Leung. "#Unrequited Love in the Cottage Industry? Managing K-pop (Transnational) Fandom in the Social Media Age." In The Korean Wave: Evolution, Fandom, and Transnationality, edited by Tae-Jin Yoon and Dal Yong Jin, 87-108. Lanham, MD: Lexington Books, 2017. Accessed September 2, 2026. https://doi.org/10.5040/9781978736214.ch-005.

Nordlund, Jan-Erik. "Media Interaction." Communication Research 5, no. 2 (1978): 150-75. Accessed September 2, 2026. https://doi.org/10.1177/009365027800500202.

Redmond, Sean. Celebrity. London: Routledge, 2018. Accessed September 2, 2026. https://doi.org/10.4324/9780203114803.

Scholz, Trebor, ed. Digital Labor: The Internet as Playground and Factory. New York: Routledge, 2012.

Silverstone, Roger, and Leslie Haddon. "Design and the Domestication of Information and Communication Technologies: Technical Change and Everyday Life." In Communication by Design: The Politics of Information and Communication Technologies, edited by Robin Mansell and Roger Silverstone, 44-74. Oxford: Oxford University Press, 1996. Accessed September 2, 2026. https://doi.org/10.1093/oso/9780198289418.003.0003.

Turner, Victor. "Liminality and Communitas." In The Performance Studies Reader, edited by Henry Bial, 79-87. London: Routledge, 2004.

Xu, Y., M. Vanden Abeele, M. Hou, and M. Antheunis. "Do Parasocial Relationships with Micro- and Mainstream Celebrities Differ? An Empirical Study Testing Four Attributes of the Parasocial Relationship." Celebrity Studies 14, no. 3 (2023): 366-86. Accessed September 2, 2026. https://doi.org/10.1080/19392397.2022.2071438.

Yin, Y. "An Emergent Algorithmic Culture: The Data-ization of Online Fandom in China." International Journal of Cultural Studies 23, no. 4 (2020): 475-92. Accessed September 2, 2026. https://doi.org/10.1177/1367877920908269.

Yin, Yiyi, and Zhuoxiao Xie. "The Bounded Embodiment of Fandom in China: Recovering Shifting Media Experiences and Fan Participation through an Oral History of Animation-Comics-Games Lovers." International Journal of Communication 12 (2018): 3317-34. Accessed September 2, 2026. https://ijoc.org/index.php/ijoc/article/view/9445.

Yoo, S. C., H. Fan, D. Piscarac, and S. Tunas Puentes. "Advertising Together for Our K-pop Idol: The Roles of Trust, Loyalty, and Perceived Value in K-pop Crowdfunding for Outdoor Advertising." Journal of Theoretical and Applied Electronic Commerce Research 20, no. 1 (2025): article 44. Accessed September 2, 2026. https://doi.org/10.3390/jtaer20010044.

Zhang, Qian, and Keith Negus. "East Asian Pop Music Idol Production and the Emergence of Data Fandom in China." International Journal of Cultural Studies 23, no. 4 (2020): 493-511. Accessed September 2, 2026. https://doi.org/10.1177/1367877920904064.

Zhou, Z., C. Su, N. Zhou, and N. Zhang. "Becoming Friends in Online Brand Communities: Evidence from China." Journal of Computer-Mediated Communication 21, no. 1 (2016): 69-86. Accessed September 2, 2026. https://doi.org/10.1111/jcc4.12147.

Websites

https://ea-cdn.eurolandir.com/press-releases-attachments./4002835/HKEX-EPS_20260318_12057241_0.PDF

https://www.isc.org.cn/article/23559.html

https://www.zhiyanbao.cn/index/partFile/1/cnnic/2021-12/1_21960.pdf

https://datacommons.org/place/country/CHN?utm_medium=explore&mprop=count&popt=Person&hl=en

https://www.cnnic.com.cn/IDR/BasicData/

An Archive of Our Own: How Fans Seized the Means of (Re)Production By Ruisi He

Abstract

This paper examines the founding of Archive of Our Own (AO3) as a pivotal moment in fan culture, representing fan's rejections of dependence on commercial platforms in order to retain creative autonomy through community owned infrastructure. Beginning with the roots of fanfiction in nineteenth-century literary retellings, to the *Star Trek* fandom of the 1960s and 70s, the paper shows how early community norms, such as collaborative creativity, exploration of sensitive themes,, and a gift economy, continued on during fandom's migration to the internet in the 1990s. The paper then examines the recurring conflicts through this migration, such as legal threats, censorship purges, and commercial exploitation. These crises exposed the problems of corporate governance, spurring AO3's creation as a donation-funded, fan-built alternative. Through its content warning system, tagging structure, and legal advocacy for transformative works, AO3 offers a model of digital self-determination that has fostered an enduring and disproportionately queer creative community. The paper concludes that AO3's history illuminates broader questions about who controls cultural narratives online, and that AO3 demonstrates communities can build and sustain alternatives to corporate-dominated digital spaces.

Introduction

In today's cultural landscape, there are few creative practices that embody the participatory nature of modern media engagement as much as fanfiction. Fanfiction is defined as "a work of fiction written by fans for other fans" (Fanlore, "Fanfiction"). Fanfiction writers take existing works—be they novels, television shows, films, or other media—and craft their own narratives using established characters, settings, and storylines. This practice extends beyond mere imitation; it functions as a dialogue between audiences and media. Fans can write stories to bring a dead character back to life, insert themselves into a narrative, or craft relationships between previously unrelated characters. Effectively, fanfiction allows fans to explore untold possibilities and challenge narrative limitations, thus "highlight[ing] the motivations and desires of readers" (Thomas 6).

Fanfiction communities have long existed at the fringes of mainstream culture. For decades, fanfiction was dismissed as derivative, inappropriate, or simply unworthy of serious consideration; even today, many of these viewpoints persist. Early critical perspectives on fanfiction often characterized it as the product of obsessive behavior or misplaced creative energy, with some scholars and critics considering it a form of cultural poaching rather than legitimate artistic expression. The predominantly female authorship of fanfiction further contributed to its marginalization, as women's creative work has historically faced greater societal scrutiny and devaluation. Yet, despite this peripheral status, fanfiction has thrived in the spaces fans have carved out for themselves—first through print zines and mail exchanges, later through digital platforms that dramatically expanded accessibility and community interaction.

The transition to digital spaces initially seemed to liberate fan communities from many of the constraints that had previously limited their reach and recognition. Early internet platforms provided unprecedented opportunities for fans to connect beyond what physical mailing could reach, share works instantly, and build diverse communities centered around shared interests. However, this digital migration came with many unforeseen problems, as fans entered new spaces and garnered more and broader-reaching attention. Questions soon arose: Who should have control over fandoms? Did they belong to the fans who created the fandoms, the authors who controlled the original franchises, or the platforms on which the fandoms were based?

In addition, fandom platforms required money, often gained through ad revenue, in order to continue running, leading to the strong influence of corporations on these spaces. The websites that fans would ultimately congregate on were subject to a variety of rules and restrictions enforced by platforms, not all of which were compatible with fan culture. Many of these issues would reach their boiling point in the mid-2000s through a series of events that exposed the fundamental incompatibility between ad-supported digital platforms and fans' creative freedoms. Sudden content purges, account suspensions for seemingly arbitrary reasons, and attempts to commercialize fanfiction created a storm that threatened the foundations of online fan communities. These crises demonstrated that fans operating on corporate-owned platforms essentially existed at the mercy of entities that rarely understood or valued their creative practices and community norms. What started as new means to spread the freedom and creativity of fanfiction became a place where fan expression could be restricted, deleted, or exploited at any moment.

It was against this backdrop that Archive of Our Own (AO3) emerged as a revolutionary response to both the cultural marginalization and digital vulnerability of fan communities. AO3 is a non-profit fanfiction archive. Its defining feature is that users can write anything, assigning almost any tag, with no censors limiting content. Despite having over eight million registered users, AO3 does not have any advertisements on its webpage. As a non-profit organization, it upholds the concept of a “gift economy,” in which users freely create content without any compensation (Cheal 1-19). Users can thus publish, read, and appreciate fanfiction without corporate interference or influence.

This paper argues that AO3 represents a milestone moment in fan culture history, when fans rejected dependence on commercial platforms and established a sustainable, community-governed creative space that operated according to historical fandom values, demonstrating how marginalized communities can develop sophisticated technological infrastructure to protect their cultural practices. More than just a website for sharing stories, AO3 embodies a radical approach to digital autonomy that challenges the conventional power dynamics between users and platforms through their fan-run, nonprofit structure and strong anti-censorship stance—creating a space where fans have total control over the work they create, despite the broader digital world becomingly increasingly commercialized.

To fully understand the rise and popularity of AO3, it is necessary to examine the origins of fanfiction as it is understood today and the nature of fandom values. Given that, the following

section provides a brief introduction to the historical origins of retelling stories and the more modern conception of fandom that rose through the advent of television and Star Trek.

The Lineage of Transformative Works and the Beginnings of Modern Fandom

Although online fanfiction did not arise until the advent of the internet, the practice of retelling stories has existed throughout human history. Long before fans gathered online to write alternate endings for their favorite TV shows, storytellers were reimagining and expanding existing narratives in creative ways, innovating and adding to works that they loved. From Dante's reinterpretation of the Bible to 19th-century penny magazines, which invited readers to submit contributions to serialized stories, people have long sought ways to participate in and shape the narratives they receive. For instance, in Louisa May Alcott's 1868 novel Little Women, the March sisters role-play as characters from Charles Dickens's The Pickwick Papers, creating poetry and stories in what they call The Pickwick Club (Alcott). Alcott thus depicts a Victorian-era practice of fanfiction.

Fanfiction began to take a form similar to what we know today around the turn of the 20th century. For example, take Arthur Conan Doyle's Sherlock Holmes novels, which completely riveted the public. As readers grew increasingly charmed by this character, they produced their own Holmes adventures, some of which were published in magazines and journals. When Doyle attempted to kill his famous detective at Reichenbach Falls, the furious public outcry forced Doyle to resurrect Holmes, with Doyle's own mother exclaiming regarding the possibility of Sherlock's death, “You won't! You can't! You mustn't!” (Panek 78). Although this is not an example of fanfiction specifically, it is evidence of the fan desire to influence the trajectory of their favorite characters, an impulse that is translated into fanfiction. This instance represents a key moment in the transition from traditional relationships of readers to literature and modern fan practices, with readers actively responding to and extending narratives to which they felt personal connections. Other examples of readers actively innovating upon works during the 20th century include The Adventure of the Two Collaborators (1923) by J. M. Barrie, a pastiche of Sherlock Holmes, and Old Friends and New Fancies: An Imaginary Sequel to the Novels of Jane Austen (1913) by Sybil G. Brinton.

It is difficult to define what constitutes fanfiction and when fanfiction practices began. Different scholars have given a variety of definitions for fanfiction. As Anna Wilson notes, “Fan fiction is a literary genre, a form, a legal category, an affective community, an aesthetic, an economy, and more, and each of its aspects demands a different approach. For this reason, attempts to create premodern fan fiction studies by first identifying a canon of premodern fan fiction texts are misguided” (Wilson). Despite this lack of consensus about the nature of historical fanfiction, modern fanfiction (and media fandom) is generally considered to have begun in the 1960s (Coppa 43). While the word had previously been used to describe amateur science fiction generally, it evolved to specifically denote fan-created stories based on existing media properties. Specifically, it was within the Star Trek fandom that the term “fanfiction” came into existence as a formal definition. The 1975 publication “Star Trek Lives!”—which explains

the history of Star Trek's fandom up to that time—is the earliest example of modern use of the word “fanfiction” cited in dictionaries. It marks the point when “fanfiction” became recognized as a distinct creative practice; moreover, it caused many fans to become aware of the existence of fanfiction and fanzines (Verba 19). As one fan noted in a 2007 issue of the zine Legacy, “This book included a chapter on fan fiction — what a revelation! You mean there were people other than me who fantasized about Star Trek stories? And actually wrote them down? I had to read these stories - now!” (Resch).

Star Trek inspired unprecedented audience engagement: as Jacqueline Guerrier notes, “Fans put their passion into action and created a new fandom, one with a size, depth, and devotion never before achieved by other fandoms before this time” (Guerrier). This was partially due to series creator Gene Roddenberry's active support of fanzines (Burns), which encouraged fans to create (Coppa 45) and thereby gave rise to one of the first substantial media fandoms. Particularly, the series' passionate fanbase established infrastructure that would define fan communities for decades to come. Star Trek fans created conventions dedicated to the show, published fanzines containing both analysis of episodes and creative works, and formed extensive mailing networks to share these materials. Star Trek fanzines like Spockanalia (first published in 1967) and T-Negative (first published in 1969) included what would now be recognized as fanfiction. As the number of fanzines available to read increased, more and more fans found out about the existence of fanzines, leading to an increased demand for more fanzines, causing a cycle that would positively reinforce the growth of the community. These early Star Trek zines established many of the tropes, formats, and community norms “which laid the groundwork for the fandoms of today” (Guerrier 28). For example, the unofficial pairing between the characters Kirk and Spock and the wide variety of content related to them helped popularize the concept of slash fiction, a genre of fanfiction that focuses on romantic/sexual relationships between characters of the same sex (Booker 126-27). In a 2000 list of popular fanfiction pairings, the fan who created the list notes how fanfiction in zines led to their discovery of slash. “New adventures, missing scenes, hurt/comfort… slash. After reading so many stories and in which [sic] Kirk and Spock cared for each other and deepened their friendship, I started to feel that something was missing. That's when I found slash” (TMar.za.net). Slash remains widely popular in fandom today, with 64 out of 100 of the most-posted relationship tags in 2024 on AO3 being M/M (male/male) (Centreoftheselights, “AO3 Ship Stats”).

Another crucial element that these early Star Trek fan works established is the predominately female authorship. It was women who created, edited, and distributed the majority of fanzines, establishing safe spaces for creative expression and often experimenting with sexuality, gender, and other marginalized topics in their written works. In addition, fanfiction provided a way for women to write when mainstream publishing was relatively inaccessible to them, especially for works on the aforementioned marginalized topics (Hellekson and Busse). These female authors shared and discussed each other's works, establishing community norms that would persist into the internet era, such as collaborative creativity and non-profit sharing of

works. Particularly, they created an environment of tolerance, which would prove hugely influential in regard to depictions of queer, non-white, or other marginalized people in the fanfiction of the time, and especially in future eras of fanfiction.

The history of fanfiction, from ancient retellings to Star Trek fanzines, shows that fanfiction isn't just some modern anomaly—it is an unbroken tradition present in storytelling from the past to the present. Literary theorists have increasingly recognized intertextuality, which is when a text references or is influenced by other texts, as fundamental to all creative production, with scholars noting that all texts exist in conversation with prior works (Alfaro). What makes contemporary fanfiction distinctive is not its derivative nature—a quality it shares with countless canonical works—but rather the communities that produce it. Unlike individual authors who are inspired by their personal feelings, fanfiction writers typically come from organized groups of media consumers who collectively transform passive consumption into active creation by reading and responding to each other's works. The community norms of the early Star Trek fandom established patterns of organization and creative engagement that would prove crucial when fanfiction communities encountered a new frontier: the internet. As fandom transitioned to early internet platforms, it carried with it both the ancient impulse to transform beloved stories and the more modern value of communal creativity; both these elements would help fanfiction resist many of the challenges that were soon to come.

The First Digital Migration

The advent of the internet in the 1990s marked a turning point for fandom—not only in how fanfiction was written and shared, but in how fans perceived their fanfiction communities. Where zines and mail-based fandom had thrived on locality, exclusivity, and analog labor, the web introduced a radically accessible, expansive model. Stories could now travel instantly across the world to be read by anyone with an internet connection. The internet thus democratized creativity, but also introduced growing pains. Tensions emerged between old and new fans, print culture and internet culture, quality control and inclusivity. Fanfiction became more popular as it became more accessible, leading to questions about profit, as well as new demand for archives. The migration online fundamentally shifted the scale and scope of fanfiction—and laid the groundwork for the issues of ownership, governance, and cultural survival that would prompt AO3's founding.

By the end of the 20th century, fan communities stood at the threshold of a massive transformation. The invention of the internet had disrupted—if not revolutionized—the print culture of fanfiction. Printed zines, conventions, and postal networks had sustained fan creativity for decades, allowing fanfiction to be delivered to paying customers. But the analog systems always had limitations. Zines required substantial financial investment to print and distribute. Geographical distance made it difficult to reach many people, and the physical nature of these exchanges constrained the scale and immediacy of fan interaction. Zines with national reach were often very selective, so most writers would write for local zines. The niche nature of fanfiction of the time also meant that fan communities did not have the subcultures they do

today, so readers did not have much of a choice about which parts of the fandom they wanted to interact with or what tropes the fanfiction available to them contained.

For many people, the internet was an accessible starting point that was much easier to understand than the complicated world of IRL zines production. The internet solved many of the aforementioned issues related to distribution and cost, helping democratize fan expression in several crucial ways. First, it dramatically lowered barriers to participation. In the past, to create a zine required capital, publishing connections, and distribution networks; now, posting a story online required only internet access and basic computer literacy. Second, it expanded fandom's reach beyond geographical and social limitations. Fans who might never have discovered zine culture—particularly international fans, younger participants, and those outside existing fandom networks—could now easily find and join fan communities online. Finally, it brought immediacy to fan interactions. Previously, it could take months (the typical publication cycle of a zine) for a fanfiction piece to reach audiences and for writers to receive comments in return; on the internet, this process took only minutes. As one fan noted in a 2002 forum discussion on zines, "I think Zines are a form of Elitism. I'm a net writer born and bred and I believe everyone has equal right to read my stories if they should choose [sic], whether they can afford to or not" (Fanlore, "Zine Elitism").

Usenet newsgroups like alt.startrek.creative ("Alt.startrek.creative") became gathering places for fans to share stories and discuss their favorite media properties. These text-based forums did not have much central moderation or commercial oversight, functioning instead as digital commons where fans established their own community norms. Email listservs which sent content through chains of emails, similarly allowed fans to form focused discussion groups, distributing stories and commentary directly to subscribers' inboxes. These early platforms lacked sophisticated interfaces but compensated with the revolutionary ability to instantly connect geographically dispersed enthusiasts who previously might never have found each other. As one fan of the show The Sentinel commented anonymously in regards to the internet on The Pitts Mailing List, "I think getting in on the ground floor of a fandom today is pretty easy. A new show comes on that you like, so you do a search on YahooGroups or the 'net in general and find the fanfic and lists, etc., and join up. I think at the birth of S&H [Starsky & Hutch] fandom, it was a lot of word-of-mouth and circulated mail, since there was no central communication system tying it all together" (qtd. in Fanlore, "Fandom and the Internet").

By the mid-1990s, web hosting services like GeoCities enabled fans to create dedicated websites for archiving fanfiction. These early archives were typically focused on single fandoms and maintained by individual fans, who manually uploaded and organized submissions. The labor-intensive nature of these efforts meant collections remained relatively small and discovery was challenging for users, as there was no centralized directory of available archives. This inability to easily reach a wider audience, combined with the financially expensive and labor-intensive process of maintaining a personal page, eventually led to GeoCities being shut down (MacKinnon).

Given these inconveniences, it was no surprise that FanFiction.net (FFN) became greatly popular when it launched in 1998. This new website allowed for much more archiving of fanfiction, becoming the first major cross-fandom archive platform. It effectively revolutionized access to fan works by creating a centralized repository where stories based on hundreds of different source materials could be found in a single location. Its category-based organization system, rudimentary search functionality, and review features transformed how fans discovered and engaged with fanfiction. For the first time, a fan of multiple properties could easily access stories across all their interests without navigating dozens of separate websites.

While FFN established the template for large-scale archives, LiveJournal (launched in 1999) expanded on the crucial social dimension of fandom. Unlike archive sites focused primarily on stories as finished products, LiveJournal also fostered community aspects of fan creativity. Fans could write personal blogs, comment on stories, and add other users as friends to discuss works in progress and new ideas, forming intimate communities. The platform's flexibility allowed for various community configurations—from large, open groups to small, invitation-only circles—mirroring the diverse social structures that had characterized analog fandom.

However, the transition of fanfiction to the internet was not without complications. With a rise in accessibility came tensions surrounding quality control, community norms, and the definition of fandom itself. Without the filtering mechanism of zine editors, online platforms featured a much wider range of fanfiction, leading to a perceived decline in quality (Kilmer) and the rise of sites such as Godawful Fanfiction ("Godawful"), which made fun of amateurish writing. The influx of new participants unfamiliar with fandom's historical norms created friction between established fans and newcomers, particularly between netfans (fans who primarily interacted online) and print fans (fans who went to conventions or subscribed to publications through mail or phone). These tensions often played out in debates over what constituted "proper" fan behavior and content—debates that would eventually contribute to the crises that necessitated AO3's creation.

Corporate Incursion and Community Crisis

As fan communities settled into digital spaces, many believed they had finally found a home—only to learn these homes were built on rented land. What many participants considered to be their creative commons were, in reality, privately owned platforms subject to corporate policies, commercial pressures, and legal vulnerabilities. Content purges, arbitrary suspensions, and outside policing from advertisers or watchdog groups soon shattered the illusion of safety and permanence. Simultaneously, ventures like FanLib attempted to commercialize fan labor without respecting its values or culture. The fundamental contradiction between fan freedom and corporate control would quickly erupt into a series of crises that threatened the stability of online fan spaces and led to a collective desire for change.

The first significant signs of these undercurrents reaching their boiling point came through increasing legal pressures on fanfiction. Specifically, copyright holders began taking

more aggressive stances against transformative works. In the 1970s, a handful of early Star Trek zines had negotiated licensing agreements with Paramount. But on the whole, fanfiction zines at the time were only distributed to small-scale local markets, insulated from corporate scrutiny by their analog nature. The internet shattered that insulation. As fanfiction migrated online, its visibility skyrocketed, and the risks of legal action targeting fanfiction increased; fans quickly became wary of this vulnerability. As one fan starkly observed in a blog post, “Fanfic is also dangerously illegal. When the stuff was distributed the old-school way, in xeroxed pamphlets at S.F. conventions, it was harder for companies to crack down on it. Now that most fanfic is on the Web, however, bored attorneys can do a Google search and come up with 16 potential lawsuits in 10 minutes” (Newitz).

In addition, several prominent authors publicly opposed fanfiction based on their works, with some sending cease-and-desist letters to archives and individual writers. Anne Rice, author of the Interview with the Vampire series, was one of those who threatened legal action; she wrote on her website in 2000, “It upsets me terribly to even think about fan fiction with my characters” (Rice). Moreover, parent groups and advertisers alike began raising concerns about inappropriate content posted on fanfiction sites. Legal threats underscored fanfiction's vulnerability; fan communities, lacking legal resources, were ill-equipped to answer key questions regarding fair use like “What constitutes a transformative work?” and “When do transformations violate copyright?”

The Harry Potter fandom, which came to prevalence in the early 2000s, became the epicenter of fanfiction copyright lawsuits. The series' unprecedented rise in popularity coincided with the internet's expansion; the number of transformative works made by Harry Potter fans was therefore explosive. This visibility made the Harry Potter fandom a prime target for corporate intervention. In what became known as the Potter Wars, Warner Bros. (WB), the studio holding the film rights to the franchise, issued a wave of cease-and-desist letters to Harry Potter fan sites in 2000–2001, demanding that they turn over the domain names of their websites or face legal action.

However, the backlash from fans was swift and organized. Unlike isolated cases where individual creators might have reluctantly complied, Harry Potter fans mobilized en masse, forming alliances like PotterWar—a coalition of fan sites that threatened a boycott of WB merchandise. This campaign, led by a 15-year-old girl named Heather Lawyer, framed WB's actions as attacks on not just individual creators but the entire fan community. As one fan site declared: “They pick on one of us, they pick on us all. We must stick together and fight this evil force together. Remember, we are stronger united than divided” (“Harry Potter Fans vs. the WB”).

The pressure worked. Faced with a public relations disaster, Warner Bros. reversed course, issuing a public apology and assuring fans that their creativity was valued. Diane Nelson, then Senior Vice-President of Warner Bros. Family Entertainment, stated: “Warner Bros. realizes how important the fans are, and in no way do we want any of them to shut down” (Buell).

Ultimately, the Harry Potter fandom's success in this legal battle represented an exception. The series' fandom was especially large, making it more effective at collective resistance. Most fan communities lacked the numbers or coordination to resist corporate pressure so effectively. While Warner Bros. retreated in this instance, the broader trend of copyright enforcement continued, shifting in later years from direct legal threats to platform-driven censorship. This continued conflict between fans and platforms soon erupted on FanFiction.net, the most popular fanfiction archive at the time.

FanFiction.net (FFN) emerged in 1998 as the first major centralized archive for fanfiction, transforming how fans accessed and shared transformative works. Its cross-fandom structure and user-friendly interface made it an indispensable hub, yet its reliance on corporate infrastructure and ad revenue ultimately exposed the vulnerabilities of commercial platforms. The first major rupture occurred in 2002, when FFN abruptly purged adult-oriented content, enforcing long-ignored Terms of Service (TOS) restrictions against explicit material. The move, likely motivated by fears of legal liability for copyright infringement by fanfiction work and advertiser backlash for inappropriate content, erased thousands of stories overnight, leaving fans scrambling to either bowdlerize their work or preserve it on smaller, niche archives. This then led to the creation of specialized archives like AdultFanFiction.net for censored works.

This deletion strongly damaged the trust of the fandom in FFN. Following the migration to the internet, the predominantly female authorship of fanfiction remained, and the experimental nature of their works became even more pronounced, especially regarding adult topics. As Sharon Cumberland notes in her essay exploring women's erotica within fan culture, “In cyberspace, where a woman cannot be criticized-or even identified-for her writing, one can see areas of curiosity and concern that could not be seen in arenas where women would have to pass through editorial hierarchies or expose themselves to the expectations of gender roles in public three dimensional life” (Cumberland). For FFN to purge all of this content was to remove a safe space for many women in fandom.

This tension between fans and platforms was further illustrated by FFN's 2005 removal of “songfics”—stories incorporating copyrighted lyrics—after record labels threatened legal action. Unlike AO3, which would later embed legal defenses for transformative works into its infrastructure, FFN had no clear mechanism to challenge such claims. Because FFN never explained its actions to fans, speculation quickly spread about FFN's motives. One fan wrote regarding the songfic purge in a timeline of major events regarding FFN in 2005, “These issues involved Cease and Desist letters sent by the MPAA [Motion Picture Association of America], alleging trademark infringement on the part of sites for using their ratings system. The other issue involved German record labels sending Cease and Desist letters to sites containing lyrics. These letters requested remuneration for every song which had the lyrics listed. While no fan fiction sites were targeted, a number of archives, including FanFiction.net, pulled all songfics in order to ensure they would not be subjected to fines” (PBworks). Whether this was truly FFN's motivation for the censorship remains unknown; what matters is that fans perceived it that way and therefore lost trust in FFN's ability to protect fans.

A decade later, history repeated itself yet again: in 2012, a user-led reporting campaign by the group Critics United triggered another wave of deletions, this time targeting Real Person Fiction (fiction written about real people, such as actors or idols), and other works that technically violated FFN's policies but had previously been tolerated. The fan backlash was swift, with users accusing the platform of hypocrisy—rules had gone unenforced for years, only to be weaponized without warning. As one LiveJournal user, simon_alexander, noted, "FFN has been a madhouse for years and blows too hot and cold for me. The admin really needs to learn to listen to their users. Obv [sic] emails and messages don't work, so walking out might" (simon_alexander). In fact, the 2012 purge would directly lead fans to migrate en masse to AO3. As AO3's Lucy Pearson observed, the "sudden and dramatic expansion" of this new platform in the early 2010s was directly tied to FFN's failures (Organization for Transformative Works, "Update on AO3 Performance Issues").

Perhaps the worst crisis that directly led to AO3's creation came in May 2007 with LiveJournal's infamous "Strikethrough" incident. With absolutely no warning, LiveJournal suspended hundreds of accounts and communities it deemed to contain "inappropriate content," marking deleted journals with a strike-through of their names. LiveJournal claimed it was conducting this purge to remove sexual content involving minors. Warriors for Innocence, a Christian watchdog organization that had taken it upon itself to identify and report "inappropriate" content on LiveJournal, claimed to have initiated this purge through their complaints. Crucially, by lobbying advertisers, Warriors for Innocence hoped to pressure LiveJournal into removing all such content. As the organization stated in a letter to the owners of LiveJournal, "If morals will not compel you to do this, then maybe loss of advertising dollars will" (Sues and Miss Y, "LiveJournal and Six Apart"). Their efforts seemed to have worked.

Warriors for Innocence claimed to be "hunting pedophiles where they fester," as per their website (Sues and Miss Y), but the execution of this censorship was both overreaching and arbitrary, catching many fanfiction pieces in its net regardless of actual content. In fact, the claim of protecting minors was likely used as an excuse by the group to target other types of content—users discovered that the purge especially targeted certain interest tags with terms like "slash" (male/male relationships) that were related to LGBTQ+ themes or adult subject matter. In addition, LiveJournal ended up purging things like rape survivor blogs, simply because they had the word "rape" connected to them (ioldanach).

The Strikethrough controversy revealed the profound vulnerability of fan communities hosted on commercial platforms. Years and years of creative work and community development were destroyed overnight, while affected users were offered little explanation or recourse for complaint. While LiveJournal eventually restored some accounts following user backlash, the incident irreparably damaged trust between the platform and its fan communities, ultimately serving as one of the greatest motivators that caused fans to move away from LiveJournal (Fiesler and Dym 11-12). To fans, it was clear that they were not the target demographic for LiveJournal, and as soon as a better alternative emerged, they would be quick to move.

In addition to these censorship crises, fans faced a new threat: commercial exploitation. In May 2007, the same month as the Strikethrough incident, a startup called FanLib launched with the aim of monetizing fanfiction. FanLib positioned itself as a professional platform for fan writers, but it failed to keep the interest of writers in mind. The company sought to profit from fan labor while offering creators minimal compensation and requiring them to sign terms of service that granted FanLib extensive rights over their work. As per their Terms of Service, "FanLib does not own any rights to Your Content. However, when you post Your Content, you are giving FanLib the right to use, reproduce, distribute and publicly display Your Content on the website or through its services (such as email notification and RSS feeds) free of charge" (FanLib.com). Specifically, FanLib hoped to sell fanfiction pieces and license them to be used in commercial productions, unlike other companies which profited primarily from ad revenue.

Additionally, the male-dominated venture capital company behind FanLib seemed to fundamentally misunderstand the predominantly female fanfiction community it sought to monetize, approaching fans as consumers to be marketed to rather than creators with established community norms and demographics. As pointed out by Henry Jenkins, one of the pioneers of fan studies, "It wasn't run by people who knew the world of fan fiction from the inside out. It was a business, pure and simple, run by a board of directors which was entirely composed of men… This last point is especially relevant when you consider that the overwhelming percentage of people who write fan fiction are women" (Jenkins). This mismatch between the men running FanLib and the women using it is evidenced by one of their ad campaigns. Ads in this FanLib campaign featured a skinny boy with the label "Life without Fan Fiction" and compared him to a muscled man with the label "Fan Fiction at Fanlib.com" (telesilla). As a fan article on LiveJournal notes, this campaign caused the mostly female community of fanfiction writers to feel "mystified and vaguely insulted. It seemed to be aimed at a stereotypical Star Trek fan with Spock ears in an ill-fitting uniform and a Klingon phrasebook" (icarusancalion). Female authors didn't want their fanfiction to be cherry picked and ranked as "the best fanfiction" by this bunch of corporate men who evidently didn't know anything about their target audience.

The response to FanLib was swift and overwhelmingly negative. Fan communities organized boycotts and created detailed critiques of the company's business model and Terms of Service. Users pointed out that FanLib was attempting to profit from fan works while offering no legal protection to creators and forcing them to assume all legal risk for their works. As LiveJournal user janecarnall noted in one such detailed critique, "There is no advantage whatsoever to a fanfic writer in registering and posting fanfic on FanLib: and while the risk of being sued may be minimal, FanLib have set it up so that by their TOS, you bear all the risk and all the costs, and are not permitted any defense against them, while they bear no risk, no legal costs - not even the legal cost of you sueing [sic] them - and take any profit that is to be had" (janecarnall).

These concurrent crises—legal threats, platform censorship, and attempted commercial exploitation—created a perfect storm that laid bare the fundamental incompatibility between corporate platform governance and fan community values. Fans faced a stark reality: the digital

spaces they had built and populated with creative work existed at the mercy of entities that didn't understand, value, or protect their communities. The convergence of Strikethrough and FanLib in May 2007 made the precariousness of fan digital existence impossible to ignore, creating both the necessity and motivation for a radical reevaluation of how fan communities should exist online.

The Birth of AO3

AO3's founding represented not merely a reaction to censorship and commercialization; it was also the manifestation of a long-held belief that fans could and should build spaces for themselves. Fan communities refused to be passive users of technology and instead became its architects. In Henry Jenkins' 1992 book Textual Poachers, he characterizes fans as active participants who appropriate elements of mass culture to create new meanings that serve their own interests and desires. Certain quotes from this book became rallying cries for fan communities, with one statement appearing on countless fan websites in the late 1990s and early 2000s: “Fan fiction is a way of the culture repairing the damage done in a system where contemporary myths are owned by corporations instead of owned by the folk” (“The Star Wars Zine Union”; Jacquez). This mindset persisted despite the issues that fandoms faced. What began as disparate expressions of frustration quickly crystallized into a movement for fan digital sovereignty—a collective determination to create infrastructure governed by and for fans. Born out of frustration but fueled by vision, Archive of Our Own was conceived as a platform where creators would not be subject to arbitrary deletion, corporate oversight, or legal precarity. Developed by fans for fans, it was grounded in the very principles that corporate sites had undermined: freedom of expression, transparency, and collective ownership.

On May 17, 2007, fan writer and coder Astolat (known professionally as Naomi Novik) posted a manifesto titled “An Archive of Our Own” to her LiveJournal; this piece is now generally considered the catalyst for the development of AO3. Written in direct response to the FanLib controversy and LiveJournal's Strikethrough, the manifesto summarized many of the complaints about existing archives that had spread through the community, and envisioned an archive “run BY fanfic readers FOR fanfic readers with no ads,” that would be “solely donation-supported with a simple and highly searchable interface and browsable quicksearch pages” (Astolat). The manifesto thereby outlined core principles that would guide the development of fan-owned platforms: open access, legal advocacy, non-commerciality, and community governance.

The fan response to this manifesto was immediate and enthusiastic. Within days, fan discussions coalesced around the practical question of how to build such an archive. Crucially, the fan community included individuals with professional expertise in relevant fields—software developers, librarians, lawyers, and nonprofit organizers—who stepped forward to volunteer their skills. As stated by Fransesca Coppa, one of the founders of AO3, “It turns out fans are everything. Fans are journalists, fans are librarians, fans are graphic designers, fans are lawyers, accountants. Fangirls do absolutely every job there is” (Charlotte).

By July 2007, these discussions had led to the founding of the Organization for Transformative Works (OTW), a nonprofit entity established to provide legal, technical, and organizational infrastructure for fan projects. By establishing a legal entity, fans created a structure that could actively advocate for their transformative works. The OTW would be able to hold assets, enter agreements, and provide institutional continuity that would stretch far beyond any individual fans. The organization's governance would be selected through elections amongst members of OTW, ensuring that the organization would remain democratic and true to fan values. To join OTW, fans simply had to donate 10 dollars—a price that was affordable to most people.

OTW's flagship project was Archive of Our Own (AO3), the fan-controlled archive envisioned in Astolat's original manifesto. Unlike commercial platforms designed primarily to maximize user engagement or data collection, AO3 was designed to prioritize creator agency in direct response to the censorship that precipitated its founding. It had no censors and no content restrictions. This founding idea—that all creative expressions deserve space, regardless of their content or perceived quality—stood in stark contrast to commercial platforms' increasingly restrictive content policies. The fact that OTW was a non-profit and supported by donations also ensured that it would not capitalize off of this archive.

AO3's philosophical commitment necessitated a technological infrastructure. The digital and legal development of AO3 proceeded deliberately, with extensive community consultation at each stage. AO3's developers made the significant decision to build the archive on Ruby on Rails, an open-source framework that aligned with the project's values of transparency and community ownership. The platform's open-source code base ensured transparency and collective ownership, allowing the fan community to maintain control over their fics. In November 2009, the archive entered closed beta testing, allowing initial users to help refine the system before broader release. By 2011, the archive had opened to all users, though its feature set continued to expand through ongoing development.

AO3 developed a comprehensive content warning system that allows creators to label any sensitive or controversial material without removing it from the archive. Along with AO3's robust filtering system, this feature allowed readers to avoid content they found objectionable without preventing others from accessing it. The system created a solution to content concerns that preserved both creator freedom and reader autonomy, rejecting the binary choice between unrestricted access and censorship that characterized many other platform policies.

In addition, AO3's tagging system was revolutionary. Rather than imposing a fixed taxonomy of categories, AO3 developed a hybrid system combining structured metadata with freeform tags. It thereby created what has been called a “curated folksonomy” (Bullard), which the OTW website describes as “a free-for-all tagging system supervised to some degree by a band of merry curators” (Organization for Transformative Works, “Spotlight on Tag Wrangling”). More formally, this classification system enabled users to classify and organize their own online content through the application of tags, allowing for a non-hierarchal and user-driven organization system. Volunteers called “tag wranglers” then connected related

concepts and linked user-created tags to a more standard form. The technical implementation of this system allowed users to create their own tags using natural language, or a language that organically evolved in a community. Through tags, creators could express what they thought readers needed to know, as well as allowing readers to find works through tags, creating a classification system that evolved with the community rather than constraining it. Highly expressive tags like “Enemies to Friends to Lovers to Enemies to Lovers Again” could exist alongside more standard categories (e.g. “angst”), accommodating both idiosyncratic expression and systematic organization. This system differed significantly from those of previous archives, which rigidly limited the tags available. A survey in 2021 noted, “All participants reported that the AO3 made it easier to find fanworks than other fannish repositories did,” with a large number mentioning “tags as a significant factor in its usability” (Black 31). Tags allowed users to find the precise type of content they wanted, and to avoid what they did not. AO3 authors “report[ed] difficulty find[ing] the right readers on other websites” (Silberstein-Bamford) partly due to problems with tagging, which would often lead readers to run into content they didn't care for.

In addition to technological infrastructure, OTW required legal expertise. To ensure AO3's autonomy, OTW needed to foster institutional protections. OTW therefore positioned fanfiction as non-infringing under fair use doctrine, developing numerous arguments to defend their position in the archive itself. AO3's Terms of Service explicitly claims works to be transformative, its submission process acknowledges transformative status, and its metadata systems clearly distinguish between original sources and transformative additions. The platform thereby embedded legal protection into the everyday functioning of the archive. By creating technical infrastructure that embodied a coherent legal theory of transformative works, AO3 challenged dominant interpretations of copyright that privileged corporate ownership over cultural participation. It offered an argument-in-code for fans' right to transform and reinterpret cultural materials. These efforts coincided with broader legal trends toward more expansive interpretations of transformative use, creating mutually reinforcing momentum for creative freedom. In another survey of fanfiction readers, “Participants mentioned being afraid that other platforms would also cave to copyright owners, but felt comfortable with AO3 as a contrast, since its policies emphasized their dedication to copyright advocacy” (Fiesler and Dym 11). By focusing on protecting the fans and upholding a strong stance regarding their principles, AO3 created trust in itself amongst the fan community, which is partially why many fans would ultimately move there.

The birth of AO3 represented a fundamental shift in how fan communities related to digital infrastructure. Rather than adapting to platforms designed for other purposes, fans created a space explicitly designed around their values and practices. Rather than accepting vulnerability to corporate decisions, they established autonomous infrastructure governed by their own community standards. This transition from dependence to sovereignty profoundly influenced how fans shared creative works and how they conceptualized digital spaces. It demonstrated that users could indeed build and maintain sophisticated technical infrastructure through collective action, expertise sharing, and volunteer commitment. As AO3 moved from concept to reality, it

established a model of digital self-determination that would prove influential far beyond fanfiction communities themselves.

Conclusion

Since 2010, fanfiction has increasingly gained recognition as a literary genre. Over the last two decades, numerous commercially successful authors have emerged from fanfiction communities. Writers like E.L. James (whose Fifty Shades series began as Twilight fanfiction) and Cassandra Clare (who wrote Harry Potter fanfiction before publishing her Shadowhunters series) have demonstrated the potential for fan writing to serve as a pathway to professional careers. By providing a platform that empowers new writers to gain a broad readership, AO3 has created a strong pathway for new writers to hone their craft. AO3 received a Hugo Award in 2019 for its efforts in preserving sci-fi and fantasy works, showing how the publishing world has grown increasingly respectful of this site and fanfiction as a whole.

Even for writers not aiming for official publication, AO3's non-profit, fan-run model ultimately challenged commercial gatekeeping that kept audiences away from writers in two crucial ways. First, it removed financial barriers to publication, allowing creators to share works regardless of their commercial viability. Second, it eliminated institutional filtering based on content, style, or subject matter, creating spaces for experimental, transgressive, or niche works that might never have found homes on commercial platforms.

By encouraging and protecting transgressive art, AO3 has facilitated a sense of community for underrepresented people. In fanfiction, authors often play with a character's racial, sexual or gender identity, refashioning mainstream narratives to better reflect underrepresented readers. As noted by Kellye Guinan in her master's thesis, these practices forge a sense of community, allowing readers to “connect to each other and share pieces of themselves that would otherwise be kept out of sight” (Guinan 17).

Fanfiction spaces have thus been particularly important for groups that are historically underrepresented in media, such as queer people and people of color. In a 2024 demographic census, an unofficial survey of 16,131 AO3 users saw 81.4% of respondents identify as LGBTQ+ (Centreoftheselights, “Survey Results”). AO3, then, has not just provided an archive of transformative fiction; it has also fostered a community for people excluded or marginalized from mainstream media. With corporations increasingly dominating online cultural production, AO3 has demonstrated that user communities can not only create but effectively sustain digital spaces that align with their values.

The history of fanfiction's migration to the internet demonstrates a recurring conflict between creative freedom and corporate control. Fans, who historically consisted of women and other marginalized writers that already lacked representation in the mainstream, struggled to free themselves from censorship, monetization, and dominant online platforms that did not share their values. Archive of Our Own is not simply a site for reading fanfiction; it is a testament to the long history of fanfiction values. Born from decades of struggle against corporate censorship,

commercial exploitation, and legal uncertainty, AO3 embodies an alternative model of online community: one built on principles of collective ownership, open access, and a gift economy.

AO3's success continues to protect fanfiction as a vital form of cultural participation that challenges dominant narratives, expands representation, and redefines who gets to shape the stories that define our world.

Today, with increasing platform monopolization, algorithmic echo chambers, and cultural homogenization on the internet more broadly, AO3 demonstrates that a different type of online community is still possible—one where users, not corporations, control creative expression. The archive's continual growth suggests that the desire for such spaces has persisted, even as commercial platforms have become the norm for most online activity.

Ultimately, the story of AO3 is not just about fans or fiction, but about power: Who gets to control cultural narratives? Who gets to participate in them? Who gets to decide what forms of creativity matter? In answering these questions, AO3 offers a hopeful vision in which the future of digital culture is written not by corporations, but by communities. Fanfiction has always been about imagining alternatives. AO3 makes those alternatives real.

Works Cited

Alcott, Louisa May. “The P.C. and P.O.” Little Women, Penguin Books, 1953.

Alfaro, María Jesús Martínez. “Intertextuality: Origins and Development of the Concept.” Atlantis, vol. 18, no. 1/2, 1996, pp. 268–85. JSTOR, www.jstor.org/stable/41054827.

“Alt.startrek.creative.” Google Groups, 2024, groups.google.com/g/alt.startrek.creative?hl=en.

Angelfire.com. “Harry Potter Fans vs. The WB.” 18 Dec. 2001. Wayback Machine, web.archive.org/web/20190806064614/http://www.angelfire.com/rpg/towerhp/goodvsevil.html.

Astolat. “An Archive of One's Own.” LiveJournal, 17 May 2007, astolat.livejournal.com/150556.html.

Black, Erin. “There's a Tag for That”: An Exploratory Study of Tag Functions in the Archive of Our Own. 2021. U of North Carolina at Chapel Hill, MA thesis, doi.org/10.17615/aagb-3n67.

Booker, M. Keith. Star Trek: A Cultural History. Rowman & Littlefield, 2018.

Buell, Ryan. “Fans Call for War; Warner Bros. Claims Misunderstanding!” Entertainment Rewired, Jan. 2001. Wayback Machine, web.archive.org/web/20060403032744/http://entertainment-rewired.com/fan_apology.htm.

Bullard, J. “Motivating Invisible Contributions.” GROUP '16: Proceedings of the 2016 ACM International Conference on Supporting Group Work, ACM, 2016, pp. 181–93, doi.org/10.1145/2957276.2957295.

Burns, Jim, editor. “Interview with (Gene) Roddenberry.” The Trek Times, no. 2, July 1976.

Centreoftheselights. “AO3 Ship Stats 2024.” Archive of Our Own, 2 Aug. 2024, archiveofourown.org/works/57864097/chapters/147284449.

Centreoftheselights. “Survey Results: Demographics.” Archive of Our Own, 23 Feb. 2024, archiveofourown.org/works/54011047/chapters/137376028.

Charlotte, host. “Fan Fiction: Origins and Cultural Impact (Since the 18th Century).” Let's Learn About…, episode 132, 7 Sept. 2022, learnaboutpod.com/ep-132/.

Cheal, David J. The Gift Economy. Routledge, 1988.

Coppa, Francesca. “A Brief History of Media Fandom.” Fan Fiction and Fan Communities in the Age of the Internet, 2006, pp. 41–59.

Cumberland, Sharon. “Private Uses of Cyberspace: Women, Desire, and Fan Culture.” The MIT Press eBooks, 20 June 2003, pp. 261–80, doi.org/10.7551/mitpress/5930.003.0020.

FanLib.com. “Terms of Service.” 31 Oct. 2007. Wayback Machine, web.archive.org/web/20071031111520/http://www.fanlib.com/termsOfUse.do.

Fanlore. “Fandom and the Internet.” Last modified 5 July 2025, fanlore.org/wiki/Fandom_and_the_Internet.

Fanlore. “Fanfiction.” Last modified 20 May 2025, fanlore.org/wiki/Fanfiction.

Fanlore. “Zine Elitism.” 2022, fanlore.org/wiki/Zine_Elitism.

"Fanfiction.net Not Showing Fics on Just In nor in Categories." 2018. Wayback Machine, web.archive.org/web/20180323184902/https://ffdotnetrants.livejournal.com/182286.html.
Fiesler, Casey, and Brianna Dym. "Moving Across Lands: Online Platform Migration in Fandom Communities." Proceedings of the ACM on Human-Computer Interaction, vol. 4, 2020, pp. 1–25, doi.org/10.1145/3392847.
"Godawful." 23 Jan. 2001. Wayback Machine, web.archive.org/web/20010123143200/http://www.go.to:80/godawful.
Guerrier, Jacqueline. Where No Fandom Has Gone Before: Exploring the Development of Fandom through Star Trek Fanzines. 2018. James Madison U, MA thesis, commons.lib.jmu.edu/master201019/580.
Guinan, Kellye Ann. Culture and Community Online: How Fanfiction Creates a Sense of Social Identity by Reshaping Popular Media. 2017. Middle Tennessee State U, MA thesis, jewlscholar.mtsu.edu/handle/mtsu/5251.
Hellekson, Karen, and Kristina Busse, editors. The Fan Fiction Studies Reader. U of Iowa P, 2014, doi.org/10.2307/j.ctt20p58d6.
icarusancalion. "Article Summing Up Fanlib." LiveJournal, 20 May 2007, icarusancalion.livejournal.com/626928.html.
ioldanach. "Support Communities." LiveJournal, 30 May 2007, innocence-jihad.livejournal.com/19200.html?format=light.
Jacquez. "Jacquez's Fanfiction." 2002. Wayback Machine, web.archive.org/web/20020211081417/http://www.dementia.org/~jacquez/writing/index.html.
janecarnall. "FanLib/Fandom: Non-Con, and Not in a Fun Way." 24 May 2007. Wayback Machine, web.archive.org/web/20160617184528/https://janecarnall.insanejournal.com/88515.html.
Jenkins, Henry. "Transforming Fan Culture into User-Generated Content: The Case of FanLib." Henry Jenkins, 22 May 2007, henryjenkins.org/2007/05/transforming_fan_culture_into.html.
Kilmer, Teresa. "Here Today, Zined Forever." The Fanfic Symposium, 23 Jan. 2001. Wayback Machine, web.archive.org/web/20010308160841/http://www.trickster.org/symposium/symp74.html.
MacKinnon, Katie. "The Death of GeoCities: Seeking Destruction and Platform Eulogies in Web Archives." Internet Histories, vol. 6, no. 1–2, 2022, pp. 237–52, doi.org/10.1080/24701475.2022.2051331.
Newitz, Annalee. "Fanfic of the Damned." Alternet, 5 June 2002, www.alternet.org/2002/06/fanfic_of_the_damned.
Organization for Transformative Works. "Spotlight on Tag Wrangling." 2025, www.transformativeworks.org/spotlight-tag-wrangling/.

Organization for Transformative Works. “Update on AO3 Performance Issues.” 2025, www.transformativeworks.org/update-ao3-performance-issues/.
Panek, LeRoy. An Introduction to the Detective Story. Bowling Green State University Popular Press, 1987.
PBworks. “FanFiction.Net History Timeline.” Oct. 2005, fichistory.pbworks.com/w/page/6133166/FanFiction.
Resch, Kathleen. “Interview with Kathleen Resch.” Legacy, edited by Jenna Sinclair, vol. 3, Beyond Dreams Press, July 2007.
Rice, Anne. “Messages to Anne Rice Fans.” 8 Apr. 2000. Wayback Machine, web.archive.org/web/20080313050018/http://www.annerice.com/readerinteraction-messagestofans.html.
Silberstein-Bamford, F. “‘Thank God for Tags’: Fanfiction as a Reading Paradigm.” New Review of Hypermedia and Multimedia, vol. 30, no. 1–2, 2024, pp. 129–47, doi.org/10.1080/13614568.2024.2369508.
simon_alexander. “A New Group to Be on the Lookout For…” LiveJournal, 9 June 2012, ffdotnetrants.livejournal.com/144727.html.
Sues, Gunny John, and Miss Y. “LiveJournal and Six Apart: It's Time to Speak.” Warriors for Innocence, 18 May 2007. Wayback Machine, web.archive.org/web/20070701191145/http://www.warriorsforinnocence.org/2007/05/livejournal-and-six-apart-its-time-to.html.
Sues, Gunny John, and Miss Y. Warriors for Innocence. 7 June 2007. Wayback Machine, web.archive.org/web/20070607070925/http://www.warriorsforinnocence.org/.
“The Star Wars Zine Union.” 16 June 2000. Wayback Machine, web.archive.org/web/20021018235336/http://www.geocities.com/zineunion/.
telesilla. “A Question for You All.” LiveJournal, 20 May 2007, telesilla.livejournal.com/557974.html.
Thomas, Bronwen. “What Is Fanfiction and Why Are People Saying Such Nice Things about It??” StoryWorlds: A Journal of Narrative Studies, vol. 3, 2011, doi.org/10.5250/storyworlds.3.2011.0001.
TMar.za.net. “Favourite Slash Couples.” 2002. Wayback Machine, web.archive.org/web/20020828071215/http://tmar.za.net/slash/couples.htm.
Verba, Joan Marie. Boldly Writing: A Trekker Fan and Zine History. 2nd ed., FTL Productions, 2003.
Wilson, Anna. “Fan Fiction and Premodern Literature: Methods and Definitions.” Transformative Works and Cultures, vol. 36, 14 Sept. 2021, doi.org/10.3983/twc.2021.2037.

A Cruel Justice: *The Justice of Southern Communities and Norms In Mark Twain's Novel The Adventure of Tom Sawyer* By Thomas Rutterer

Introduction

Oxford dictionary defines justice as "maintenance of what is just or right…" This definition, however, is wrong as justice is the "the product of regimes and norms" in a community (Heinze 1). This look on justice reflects Mark Twain's use of cruel irony in the novel *The Adventures of Tom Sawyer's* main antagonist's, I-word Joe[1], to convey the injustice of racism. Joe's treatment reflects his hatred and vindictiveness, pleading, death, and Tom's town's, St. Petersburg's, true meaning of justice. Social and community norms of the 1840s South of racism cause the unjust treatment of Joe as a child, which Mark Twain uses to bold southern racism and its consequences of revenge.

Preview

This paper will highlight three key moments of Joe's murder, Joe's pleas of murderous intent, and Joe's death scene that will emphasize racism in Tom Sawyer's town. Each moment will convey Joe's ostracization by his community, Joe's wicked treatment when he was a child, and how even the highest in the St. Petersburg are against him, which will demonstrate racism and racial stereotypes in southern society. Additionally, to expand the conversation on *The Adventures of Tom Sawyer*, this paper will make a comparison between *The Concept of Injustice* by Eric Heinze and the demonstration of racism from I-Joe's life which will provide another way to appreciate the cruel irony of I-Joe. Twain intends to convey the true meaning of justice and heroism, and that heroes and villains are not as different as they are perceived to be.

Revenge

In Mark Twain's *Adventures of Tom Sawyer,* I-Joe, who can be considered the villain of the story, is shaped by the treatment he faces in his early ages. The treatment Joe faces is based on his racial profile. For this treatment Joe lashes out to help himself feel better and then faces the consequences for his pleading. I-word Joe's first plea for respect from his colleagues is in the act of revenge. This act of revenge takes place in the graveyard as Joe, Dr. Robinson, and Muff Potter, one of the town drunks, sneak into the yard to rob a body for the purposes of studying the anatomy of it. Dr. Robinson, looking at the body, tells Joe how he has been paid in advance. Joe replies: "Yes and you've done that more than once,"(Twain 65). Joe continues his reply, '"Five years ago [, Dr Robinson,] you drove me, a way from your father's kitchen…I swore I'd get even with you if it took a hundred years…when I [came] to ask for some food,"'(Twain 65). Joe then proceeds to kill the doctor for his revenge. Twain uses I-Joe's circumstances and backstory of injustice to demonstrate the flaws of racism and the consequences. Because of Joe's backstory, this act of revenge demonstrates that Joe is not just a flat character of a murderer, but instead a

[1] The word indicated by calling Joe "I-word Joe" is to avoid the slur used in the original text

pleading soul who is pillaging for a chance at treasure, the treasure of respect. Additionally, the word "'drove"' demonstrates racism in the example of denying food to the man. According to Etymonline, the word "drove" is the past tense of drive, which comes from the definition of hunting or pursuing deer, symbolizing Dr. Robinson's thought of Joe as an animal to shoo away.

After the threat by Joe, Dr. Robinson, "struck out suddenly," demonstrating the first blow the argument was Robinson's act of driving away Joe after the fact Muff Potter drops "his knife," symbolizing his peacefulness not wanting to fight (Twain 65). The first strike being from Robinson illustrates that Joe is the victim of societal injustice as though he had threatened the man he had not taken a strike at him yet, conveying the white man as the instigator. Glenn Loury illustrates the point of racial injustice and it's consequences perfectly with: "An important consequence of racial stigma is "vicious circles" of cumulative causation,"(Loury 2). The "vicious circle" in Joe's case is the cycle of "being hurt and hurting to be" where Joe has been attacked psychologically by the denial that he must seek mental closure, dealing what has been done onto him to another physically. In conclusion, Twain uses Joe's circumstances and murder as a way to illustrate the consequences of racial inequality in southern communities by matching Joe to the common vicious cycles of racism and cruelty.

Desire

Twain's cruel irony is additionally demonstrated in Joe's pleas to his colleagues and desire for respect. After Joe has been accused of murder, he disappears until he decides to attack Widow Douglas' house. The attack is because Joe thought of Mr. Douglas as: "'the justice of the peace that judged me for a vagrant,'"(Twain 176). Joe describes this judging as "'being rough with me…It's a millionth part of it,"' Joe continues, "'He had me horsewhipped…like an [n-word],'"(Twain 176). Joe describes his whipping which helps convey the message of racism in Tom's community because whipping is a common practice to punish slaves. The message is illustrated by the word "'like a [n - word],"' as Joe, who is discriminated against, looks beneath himself at black people demonstrating Joe as a marionette of cruel irony to exemplify the wrong in racism. The whipping of Joe which uses a horsewhip is symbolic of the racism in Tom's community as Joe is being compared to a horse. Horses are work animals used for the gain of others, demonstrating how Joe is a horse being worked and discriminated against by Tom's community because of his race.

Joe does have compassion for those who have not wronged him for instance: "'I would kill him if he was here but not [Widow Douglas],"'(Twain 176). This compassion for others conveys that Joe is not just a lunatic who hates the world but a vigilante seeking justice for himself. After hearing this murderous talk, Huck Finn, who was eavesdropping after he saw Joe leave his room at the tavern as he already knew that Joe had committed murder before, runs to the Tavern where he is greeted by a Welshman who mobilizes and runs off to attack Joe. Huck in this sense, can be called a hero, even though he grew up similar to Joe he did the right thing in the moment. Huck and Joe reflect to the knight and the dragon in Evan Poole's Article "Heroes vs. Villains" the article states: "[The hero and the villain] both kill for their own pleasure, the

monster gets a thrill out of murder and the knight feels satisfaction after slaying the monster. Both struggled hard to achieve their goals. Both have legitimate, understandable reasons for doing what they do,"(Poole 1). Huck runs to the Welshmen to kill Joe while Joe wants to attack Widow Douglas both have tried to kill. Poole says: "The justifications for the hero's actions are just as valid as the villain's, so why do we treat the hero differently,"(Poole 1). The difference for Joe and Huck is race, one white, the other a native american "half breed," demonstrating that Joe is used by Twain to convey racism in southern society from the outcome of Huck and Joe. In conclusion, Joe's pleas of murderous intent and how he is treated differently from Huck Finn demonstrate the cruel irony of Joe's character and the ideals of not condoning racism relevant.

Joe's Expiration

Joe's final call for respect, demonstrates his character being meant for cruel irony is his death. Tom, after finding a way out of the cave he got lost in after running into it past his friends he ends up lost with Becky Thatcher, discovers the Thatchers, the rich and powerful members of Tom's society, constructed an iron door to seal off the cave because Tom and Becky got lost in the cave. Tom, after hearing this information, runs to the cave and has the door opened because Tom, while he was lost in the cave, saw Joe walking around in the cave from which Joe would probably leave from the known entrance. Tom sees Joe's "longing eyes…fixed" on the "light and the cheer of the free world" at the opening (Twain 199). Tom then takes a closer look at Joe which the narrator states: "Joe's bowie knife lay close by its blade broken…[the stone pillar] had been chipped and hacked…the only damage was to the knife"(Twain 200). The bowie knife's broken blade signifies the strikes and pleas of Joe and that they have come to nothing and that he will *always* be characterized as his race by others around him. This demonstrates Twain's cruel irony fixed into Joe's character that no matter how hard one tries they cannot change other's feelings around them. People can only change the feelings they have about themselves. The bowie knife connects to Muff Potter's knife in the first instance of Joe's revenge, his murder in the graveyard, where Muff Potter drops his knife after he sees the fighting of Robinson and Joe begin; additionally, in this instance, Joe drops the knife after his death, symbolizing his plea to make the world stop like Muff Potter did and that pacifism seen by Muff in the graveyard.

The final demonstration of Joe's desperation to leave his prison. Joe has been put into the prison by the upper-class men of the society. In this prison: "ordinarily one could find a half dozen bits of candle." However, as the narrator points out, "there were none"(Ibid). The narrator states that there are no candles left, illustrating the fact that Joe ate all of the candle bits just to live a little longer, to hope society would finally shine on him, but his moment never comes. The point both Muff and Joe are trying, with the dropping of the knives, to make is one that translates to the belief of nihilism, nothing really matters, and as Princeton University Press puts it: "the attempt to reorganize life according to the patterns from which we have departed when we fell into the crisis of nihilism does not lead to our liberation from it,"(Princeton University Press 6). Joe, in the act of dropping his knife, is trying to put the world back to where he "departed [from] when [he] fell" to the act of revenge as in the end nothing really mattered, the revenge or the

pleas, all that amounted to was his own death. In conclusion, Joe's final act of dropping his knife conveys Twain's cruel irony, demonstrating that even if the victim of racism struggles and winces, if the community is against him, nothing in the end is amounted to, except the victim's own sorrow.

Inherent Justice

The final element of Twain's cruel irony is in the aspect of what Joe life means to the reader, which is how justice is shaped; additionally, the shape justice reflects I-Joe, his life, and his cruel irony, demonstrating how justice is corrupt in a sense. For example, Eric Heinze, co-founder and director at Queen Mary University of London Law School, dives deeper into the theory of cruel justice. Heinze reflects the words of Plato, a Greek philosopher, writing injustice as not the opposite of justice but the "product of regimes and norms of justice"(Heinze 1). Heinze extends his claim to who shapes justice on victims: "Law and justice are risible weapon, forged by mediocre, cowardly multitude, the weak and the meak, who, at the hands of the powerful, merit not justice but disdain,"(Heinze 5). Justice in Tom's community is shaped by the white people enforcing "justice" on the other races and, as demonstrated by the murder of Robinson, Joe is powerful enough to kill the weak. The weak in this case are the white people who use polluted tactics of not giving food or whipping for asking for it. The nuance of the rules of "justice" being made by the weak exemplifies that the weak will use anything to be powerful because of the menial want of the weak or unfilled peoples' want for power. The extent to which this power hungry nature goes is to the extent of turning justice, what should be the right thing to do, into a system from which their race can prevail. In other words, turning justice into a rigged survival of the fittest.

Tom's community are thieves, taking from Joe and his future to amass their power over him, as Heinze says, "the select few, the clever and the daring, ought not to toil when they can prosper by theft and stealth"(Heinze 5). Tom's community prospers from the wealth they gain from taking from Joe which demonstrates the society's bias of white over other races. Heinze's concept of justice matches with Joe's lack of the justice system. Joe, after committing murder in the graveyard, blames the murder on Muff Potter because Muff's knife was the weapon. Muff brought to court is saved by Tom Sawyer's confession of what happened at the graveyard as he had been at the graveyard when the murder occurred. After Tom confesses, Joe, "quick as lighting…sprang…[and] tore his way through all opposers"(Twain 147). Heinze's concept of justice matches Joe's quick run out of the courtroom as Joe even knowing that a trial can be a wrong decision. Another match to Heinze's article is the way the people are described, "opposers," demonstrating how everyone is against Joe. Joe in this way runs out of a room of hatred not just for his actions, but also the town's normal hate of non-white people. In summary, Heinze's *Concept of Injustice* matches Twain's use of cruel irony in I-Joe for the illumination of racial bias because of the white of Tom's community who take away Joe's childhood and "good" years to benefit their own lives with "theft", justice is not just what is right but instead what is shaped by the society. Therefore, justice, on these terms, is always unjust to someone.

Conclusion

Finally, Joe's revenge, pleas, and death, illustrate the true meaning of justice in Southern society. These moments represent Twain's cruel irony reflected in I-Joe used to demonstrate the racism in Tom's and other communities and how it can ruin someone's life. Racism is still a key issue today and can ruin a person's day and be remembered through their life. As Blanche DuBois says in the movie, "A Streetcar Named Desire," we "'rely on the kindness of strangers,"' for justice is not a just or right thing, instead justice is shaped by the community and its "norms."

Works Cited

Heinze, Eric. “The Concept of Injustice.” SSRN eLibrary, 2012, papers.ssrn.com › sol3 › papers.

Loury, Glenn. “Focus University of Wisconsin–Madison Institute for Research on Poverty.” Wisconsin University, 2005, irp.wisc.edu/publications/focus/pdfs/foc241a.pdf.

Twain, Mark. The Adventures of Tom Sawyer. Bantam Books, 1981.

Princeton University Press. “I Nihilism Curiosity about the World, as We Know, Sometimes Leads to Philoso.” Princeton University Press, assets.press.princeton.edu/chapters/s9593.pdf. Accessed 29 Oct. 2025.

Poole, Evan A. (2017) "Heroes vs. Villains," Sierpinski’s Square: Vol. 2 : Iss. 1 , Article 9. Available at: http://digitalcommons.wku.edu/siersquare/vol2/iss1/9

Warner Bros., Warner Home Video. (1951b). A streetcar named Desire. United States, United States.

Memory, Belonging, and the Gradual Normalization of Exclusion: A Case Study of Holocaust Survivor Gabriella Karin By Honor Chacon

Abstract

Holocaust survivor testimony is both a record of individual experience and a source for examining the social processes that make persecution possible. This paper analyzes the testimony of Gabriella Karin, a Holocaust survivor born in Bratislava, Czechoslovakia, in 1930, using a personal interview as its primary source and historical scholarship and institutional archives as secondary sources. Karin's experience demonstrates that the Holocaust did not begin with mass killing but with a gradual normalization of exclusion: legal restrictions, economic dispossession, social isolation, and the redefinition of Jews as outsiders. Her story also complicates the assumption that assimilation guarantees belonging, since her secular and integrated family remained vulnerable once antisemitism became state policy. Finally, the paper examines survivor testimony as a form of historical evidence and considers how Holocaust memory shapes contemporary understandings of antisemitism, immigration, and Zionism. The paper argues that Karin's testimony is most valuable not because it provides a simple analogy for present-day politics, but because it reveals recurring mechanisms of dehumanization and the moral significance of individual choices made within oppressive systems.

Keywords: Holocaust testimony, antisemitism, Slovakia, assimilation, immigration, historical memory, Zionism

Introduction

The Holocaust is often remembered through its most recognizable endpoints: ghettos, deportation trains, concentration camps, mass shootings, and killing centers. Those images are essential to understanding the scale and brutality of Nazi genocide, but they can also make the Holocaust appear as though it emerged suddenly and fully formed. The testimony of Holocaust survivor Gabriella Karin complicates that impression. Her memories of childhood in Bratislava show a society in which ordinary life, assimilation, and neighborly familiarity existed before legal exclusion and fear became normalized. Her survival story therefore raises a question that extends beyond biography: how does a society move from everyday coexistence to the systematic persecution of a minority?

Karin was born in Bratislava in 1930, when the city was part of Czechoslovakia. In a personal interview, she described a childhood that initially felt secure and ordinary. Her parents owned a delicatessen, and the family was largely secular and integrated into local life (Karin). Holocaust Museum LA similarly identifies Karin as a Bratislava-born survivor whose parents owned a delicatessen and later secured false identity papers for her when anti-Jewish persecution intensified ("Survivor Talk"). These details matter because they place the Holocaust within an ordinary social world rather than outside it. Karin did not begin life in hiding. She became someone who had to hide because the political and social meaning of being Jewish changed around her.

This paper uses Karin's interview as a primary source and places it alongside historical material from the United States Holocaust Memorial Museum, Yad Vashem, Holocaust Museum LA, the Museum of Tolerance, Oxford scholarship on survivor testimony, Pew Research Center, and United Nations historical material. Rather than treating one survivor's memory as a complete account of the Holocaust, the paper uses Karin's experiences as a case study through which larger themes can be examined. Survivor testimony provides access to emotions, perceptions, and everyday experiences that administrative records often cannot capture, while documentary sources help establish the broader historical context in which those memories took place (Lipsitz; United States Holocaust Memorial Museum, "Historian Peter Black").

Karin's testimony supports three related arguments. First, persecution becomes especially dangerous when exclusion is introduced incrementally and accepted as normal. Second, assimilation does not necessarily protect minorities when political institutions redefine national belonging in racial, ethnic, or religious terms. Third, survivor testimony remains important because it connects large historical processes to individual lives and forces later generations to confront the moral choices made by ordinary people. These lessons can inform contemporary discussions of antisemitism, immigration, and Zionism, but only if historical comparison is used carefully. The purpose of studying Karin's story is not to claim that every modern conflict resembles Nazi Europe. It is to identify how prejudice, legal inequality, silence, rescue, displacement, and memory operate within human societies.

Method and the Value of Survivor Testimony

The central primary source for this paper is a personal interview with Gabriella Karin conducted after hearing her speak at the Museum of Tolerance in Los Angeles. The interview covered her childhood, wartime experiences, hiding, immigration, identity, antisemitism, Zionism, and the responsibilities she associates with historical memory (Karin). Because the interview occurred many decades after the events described, it is important to treat it as oral history rather than as a transcript of events recorded at the time. Oral testimony is shaped by memory, later experience, and the questions an interviewer chooses to ask.

That limitation does not make testimony historically unimportant. Neal Lipsitz explains that Holocaust testimony has been used both to deepen the historical record and to encourage moral reflection, and that testimony takes many forms, including structured interviews, written narratives, and audiovisual accounts (Lipsitz). The United States Holocaust Memorial Museum maintains tens of thousands of oral histories because eyewitness testimony preserves dimensions of experience that other sources cannot. Former USHMM senior historian Peter Black emphasizes that survivor testimony can communicate what life under persecution felt like in ways that official documentation cannot, even while historians must remain aware of the normal imperfections of human memory (United States Holocaust Memorial Museum, "Historian Peter Black").

For that reason, this paper does not ask Karin's testimony to prove every historical fact on its own. When possible, biographical and historical details are checked against museum records and Holocaust archives. For example, institutional accounts confirm that Karin was born in

Bratislava in 1930, was provided false papers, spent time in a convent, and later hid with seven other Jews in the apartment of Karol Blanar, a non-Jewish lawyer who risked his life to shelter them ("Survivor Talk"; Museum of Tolerance, "Gabriella Karin"). Yad Vashem's list of Righteous Among the Nations for Slovakia records Karol Blanar together with his parents, Vincent and Katarina Blanar, as recognized rescuers (Yad Vashem, "Righteous Among the Nations"). These records allow Karin's personal memories to be placed within a verifiable historical framework.

This approach also helps preserve an important distinction between testimony and generalization. Karin can speak authoritatively about what she remembers, what she felt, and what she observed. Her experience can illustrate broader historical patterns, but one survivor cannot represent every Jewish person in Slovakia or every Holocaust survivor. The strength of the case-study method is therefore not that it turns one life into a universal story. It is that it makes large historical processes visible at the scale of a human life.

From Discrimination to Persecution in Wartime Slovakia

Karin's childhood unfolded during a period of rapid political transformation in Central Europe. The independent Slovak state created in 1939 allied itself with Nazi Germany and enacted increasingly severe anti-Jewish measures. The United States Holocaust Memorial Museum notes that between 1939 and 1941 Slovakia adopted a series of antisemitic racial laws that became known collectively as the Jewish Code. In 1942, Slovakia became the first Axis ally to agree to deport its Jewish population to German-occupied Poland (United States Holocaust Memorial Museum, "Axis Powers"). Yad Vashem records that deportations from Bratislava accelerated in 1942, with Jewish men, women, and families sent through Slovak camps and onward to Majdanek and Auschwitz (Yad Vashem, "Deportations").

The Jewish Code is particularly important to the argument of this paper because it illustrates how persecution can be institutionalized before mass killing begins. The law restricted Jewish participation in public, economic, and social life and translated antisemitic ideology into ordinary rules governing work, property, movement, education, and identity (United States Holocaust Memorial Museum, "Axis Powers"). The danger of such laws did not lie only in the immediate hardship they created. They also trained the broader society to accept the idea that Jewish citizens could legitimately possess fewer rights than their neighbors.

Karin's recollection of her early childhood makes this transformation vivid. She remembered her family as assimilated and her prewar life as relatively normal. In her account, antisemitism became frightening not because every neighbor instantly became violent, but because the environment around her changed step by step (Karin). People adapted to new restrictions. Behaviors that would once have seemed extraordinary became routine. The change in social expectations mattered alongside the change in law.

This process helps explain why the Holocaust cannot be understood only through its final stage. Genocide required institutions capable of identifying, isolating, dispossessing, transporting, and murdering people on a massive scale, but those institutions operated within societies that had already been taught to distinguish between people whose rights mattered and

people whose rights could be withdrawn. In Slovakia, antisemitic law formalized that distinction. By 1942, deportation was not an isolated eruption of violence; it followed years of escalating exclusion.

Karin's experience therefore supports a more precise understanding of the phrase "warning signs." Warning signs are not evidence that every discriminatory policy will inevitably end in genocide. History does not operate through automatic repetition. Rather, the Slovak case demonstrates that the normalization of unequal citizenship can make more extreme actions politically and socially easier to carry out. When a targeted population is repeatedly described as dangerous, disloyal, alien, or undeserving, the moral barrier against harming that population can weaken.

The scale of what followed in Slovakia makes the progression unmistakable. USHMM reports that approximately 57,000 Jews were deported from Slovakia between March and October 1942, and that later deportations and killings followed after German forces occupied Slovakia in 1944 (United States Holocaust Memorial Museum, "Axis Powers"). Yad Vashem's account of Bratislava describes arrests, temporary detention sites, forced transfers, and deportations to killing centers (Yad Vashem, "Deportations"). These numbers reveal the magnitude of the crime. Karin's testimony reveals what that transformation meant to a child who had once believed her family belonged.

Hiding, Rescue, and Individual Moral Choice

The destruction of legal protections forced Karin's family to rely on deception, secrecy, and the willingness of non-Jewish individuals to help them. Holocaust Museum LA reports that her parents obtained false papers identifying her as Christian and sent her to a convent in Bratislava. Karin later hid with seven other Jews in the apartment of Karol Blanar, a young lawyer who knowingly risked his life to shelter them ("Survivor Talk"). The Museum of Tolerance likewise identifies Karin as a survivor who hid in a convent and later in Blanar's Bratislava apartment (Museum of Tolerance, "Gabriella Karin").

The existence of false papers demonstrates how completely citizenship had failed to protect Jewish people. Survival could depend on being able to present oneself as someone else. Identity documents that should have confirmed legal personhood instead became tools through which the state marked people for exclusion. The false Christian identity assigned to Karin was therefore more than a disguise. It represented a reversal of the normal relationship between a person and the law: the truth of who she was made her unsafe, while a fabricated identity increased her chance of survival.

At the same time, Karin's survival depended on people who rejected the moral logic of the regime. Blanar's actions are especially important because they show the difference between political power and moral agency. He could not end the Holocaust. He could not repeal the Jewish Code or defeat Nazi Germany. He could, however, decide whether the people in front of him would be abandoned. By sheltering eight Jews, he transformed private space into a site of resistance.

Karin's testimony repeatedly returns to the importance of people who helped rather than harmed. That emphasis complicates accounts of the Holocaust that divide the world too neatly into all-powerful perpetrators and completely powerless bystanders. The range of choices available to individuals varied dramatically, and resistance could carry mortal risk. Yet the fact that rescue was dangerous is precisely what gives it moral significance. The nuns who sheltered Karin and the Blanar family who assisted Jews did not act because the law protected their decisions. They acted in opposition to the system surrounding them.

This point matters for the paper's broader argument about normalization. Social norms are reinforced when people comply with them, but they are also challenged when people refuse. The same society that produced legal exclusion also contained individuals willing to violate those expectations. Karin's story therefore demonstrates two opposite forms of ordinary human behavior: adaptation to injustice and resistance to it. The Holocaust was created by institutions and perpetrators, but its history also includes the choices of neighbors, clergy, professionals, and families who decided whether to participate, remain silent, or intervene.

Karin later worked to preserve the memory of those choices. The Museum of Tolerance has included her among survivors who have shared testimony for decades, and Holocaust Museum LA documents her continuing role as a public speaker and educator (Museum of Tolerance, "Witness to Truth"; "Survivor Talk"). Her postwar work shifts the meaning of survival from a private outcome to a public responsibility. She survived because others refused indifference; she now uses that survival to ask later generations what they will do when they encounter dehumanization.

Assimilation and the Fragility of Belonging

One of the most intellectually important aspects of Karin's testimony is the challenge it poses to a common assumption about assimilation. Her family did not describe itself as socially isolated from Slovak society. Her parents owned a business, the family was largely secular, and Karin remembered herself as integrated into the community around her (Karin). Yet these characteristics did not prevent the state from defining the family as Jewish and subjecting it to antisemitic restrictions.

Assimilation is often discussed as though belonging is something minorities can earn through language, work, education, cultural participation, or loyalty. Karin's experience reveals the limits of that idea. A person may feel fully part of a country and still be excluded when the dominant political system changes the criteria for membership. Once racial or religious ancestry becomes more important than civic participation, assimilation can become irrelevant.

This is one reason the history of antisemitism is difficult to reduce to simple explanations about cultural differences. European Jews occupied many different positions: religious and secular, wealthy and poor, culturally distinct and highly assimilated. Nazi racial ideology targeted people based not simply on religious practice but on ancestry. In Slovakia, the Jewish Code formalized categories that restricted the rights of people defined as Jews regardless of how integrated they may have felt in local society (United States Holocaust Memorial Museum, "Axis Powers").

Karin's story therefore reframes the concept of belonging. Belonging is not only a feeling held by the individual. It is also a relationship between the individual and political institutions. A person may understand a country as home while the state withdraws recognition of that person as an equal member. When that happens, citizenship becomes fragile because the protections attached to it depend on whether institutions continue to recognize the person as part of the national community.

This insight has relevance beyond the Holocaust, but comparison must remain disciplined. Modern democratic debates over immigration, citizenship, border enforcement, or national identity are not automatically equivalent to Nazi persecution. Equating them carelessly can distort both the Holocaust and present-day politics. The more useful lesson is narrower: a society should distinguish between debating policy and dehumanizing people. Political communities will always make rules about citizenship and immigration. The danger appears when entire populations are described as inherently threatening, morally inferior, or undeserving of basic dignity.

Karin's experience demonstrates why that distinction matters. Her family's assimilation did not protect them because antisemitic politics redefined them collectively. The individual disappeared behind the category. Once that happened, the question was no longer what Karin's family had contributed to Bratislava or how deeply they identified with their home. Their assigned group identity became more important than their actual lives. This is one of the central mechanisms of dehumanization: the replacement of individual human beings with a stereotype about a group.

Immigration, Home, and the Author's Family History

The interview also changed how I understood my own family's history. My grandparents immigrated to the United States from Michoacán, Mexico, during the late 1960s. Their experience was fundamentally different from Karin's. They were not escaping a program of genocide, and placing the two stories on the same moral or historical scale would be inaccurate. The comparison is useful only at a more limited level: both stories raise questions about what "home" means, why people leave familiar places, and what it means to seek belonging elsewhere.

The timing of my grandparents' migration also places their story within a major transition in American immigration. Changes to U.S. immigration law in 1965 replaced the national-origins quota system with a structure that placed greater emphasis on family reunification and employment. The law and the end of the Bracero program helped reshape migration between Mexico and the United States in the decades that followed (Pew Research Center, "Nation's Immigration Laws"; Passel et al.). My family's story belongs to that much larger history of post-1965 migration.

Growing up, I heard about the difficulty of adapting to a new language, new customs, and new expectations. My grandparents worked hard while trying to create opportunities for their children and grandchildren. Their sacrifices are among the reasons I am able to pursue an education today. Karin's story made me pay greater attention to the emotional dimension of

migration: leaving home means giving up familiarity even when the destination promises safety or opportunity.

At the same time, comparison has limits. Karin's home became dangerous because a state and its allies targeted Jews for persecution and destruction. My grandparents left rural Mexico in search of greater opportunity. These are not interchangeable experiences. Treating them as interchangeable would weaken rather than strengthen the argument. What connects them is not equal suffering but the human importance of belonging, security, and hope.

This distinction also changes how I think about the word "American." If national belonging is defined only by ancestry or birthplace, immigrant families can remain permanent outsiders regardless of how much they contribute. If it is understood civically, national identity can include people who arrive from elsewhere and participate in the country's institutions, communities, and public life. Karin's history shows what can happen when ancestry becomes the decisive test of belonging. My grandparents' history shows why the ability to build a new home can matter across generations.

Historical Memory and Contemporary Antisemitism

Karin's public work as a survivor-educator demonstrates that Holocaust testimony is not only about reconstructing the past. It is also about deciding what later generations will remember. The Museum of Tolerance's "Witness to Truth" project includes Karin among survivors whose testimonies have shaped the museum's educational work, and the institution notes that tens of thousands of visitors hear survivor and second-generation speakers each year (Museum of Tolerance, "Witness to Truth"). Such programs reflect a larger shift in Holocaust remembrance: as the survivor generation becomes smaller, institutions increasingly preserve recorded testimony, archives, art, and educational programs so that personal memory can outlast the people who experienced the events directly.

The urgency of that work is connected to the persistence of antisemitism. Contemporary antisemitism should not be described as though the United States or other democracies are simply repeating 1930s Europe. The historical conditions are different, and responsible comparison requires attention to those differences. Yet anti-Jewish prejudice remains measurable. In a 2025 Pew Research Center survey, substantial majorities of both Democrats and Republicans said Jews in the United States experience at least some discrimination (Pew Research Center, "How Much Discrimination"). The significance of this finding is not that public perception proves every allegation of antisemitism, but that prejudice against Jews remains visible enough to be widely recognized across partisan lines.

Karin's testimony helps explain why the form of prejudice matters as much as its intensity. Antisemitism is not limited to physical violence. Historically, it has operated through conspiracy theories, assumptions of collective disloyalty, economic stereotypes, religious hatred, political scapegoating, and claims that Jews do not truly belong to the societies in which they live. The Holocaust represents the most catastrophic outcome of antisemitism, but the ideology did not begin with extermination. It developed through older traditions of exclusion that Nazi politics radicalized and converted into state policy.

This is why memory must be more than ceremonial. Remembering six million murdered Jews is essential, but historical memory also requires studying the processes that made their persecution possible. If the Holocaust is remembered only as an unimaginable evil committed by uniquely monstrous people, it becomes easier to distance the present from the past. Karin's story is unsettling precisely because it begins with recognizable things: a family business, school, neighbors, paperwork, laws, housing, and the question of whether another person will help. Survivor testimony also prevents historical victims from becoming anonymous. Statistics communicate scale, but scale can be psychologically difficult to comprehend. One child hiding under a false identity can reveal the human consequences of a policy more directly than a number in a table. This does not mean emotion should replace evidence. It means that testimony and documentation answer different historical questions. Deportation records can show how many people were removed. Testimony can show what it meant to live in anticipation of removal.

Karin's continuing emphasis on compassion is therefore historically significant rather than merely inspirational. Her story does not deny the scale of evil; it asks what forms of behavior remained possible within it. That question shifts Holocaust memory from passive remembrance to ethical responsibility. The point is not to imagine that an individual can always stop a violent political system. It is to recognize that systems are experienced through countless human decisions about whether to obey, assist, conceal, report, protect, exclude, or remain silent.

Zionism, Jewish Security, and Competing Histories of Displacement

One of the most challenging parts of my conversation with Karin concerned Zionism and the Israeli-Palestinian conflict. Before speaking with her, I tended to encounter Zionism primarily through current political arguments. Karin's life pushed me to consider the movement within a longer history of Jewish insecurity in Europe. That historical context does not settle present political disputes, but it helps explain why the idea of Jewish self-determination became deeply meaningful to many Jews, especially in the aftermath of the Holocaust.

Modern political Zionism developed before the Holocaust, during the late nineteenth century, in response to a combination of Jewish historical ties to the land of Israel, European nationalism, and persistent antisemitism. The United States Holocaust Memorial Museum explains that Zionism was a modern political movement advocating Jewish autonomy and that rising antisemitism was one of the forces that shaped its development (United States Holocaust Memorial Museum, "Holocaust Survivors"). The Holocaust then transformed the urgency of the question. After the war, hundreds of thousands of Jewish survivors were displaced, and debates over immigration, refuge, and a Jewish state became inseparable from the recent failure of European states to protect Jewish populations.

Karin's biography makes that history personal. Holocaust Museum LA reports that after the war she moved to Israel and later to Los Angeles ("Survivor Talk"). For a survivor whose citizenship had failed to protect her as a child, the idea of a place in which Jews possessed political self-determination could not be only an abstract theory. It was connected to the memory of dependence on governments that had either persecuted Jews or failed to save them.

Understanding that perspective, however, does not require ignoring Palestinian history. The creation of Israel and the 1948 Arab-Israeli war were also accompanied by large-scale Palestinian displacement. United Nations historical materials describe the emergence of the Palestinian refugee crisis during the 1947-49 conflict and the displacement of hundreds of thousands of Palestinians (United Nations Information System). Any serious historical account must therefore recognize that Jewish and Palestinian collective memories are shaped by different but deeply consequential experiences of insecurity and displacement.

This is where historical empathy becomes difficult but necessary. Political arguments often reward narratives in which only one group's suffering is considered legitimate. Karin's testimony led me toward the opposite conclusion. Understanding why the Holocaust intensified Jewish support for a secure national homeland does not require endorsing every policy of every Israeli government. Recognizing Palestinian displacement and aspirations for self-determination does not require denying Jewish historical trauma or treating antisemitism as acceptable political speech. These positions can coexist because historical understanding is not a zero-sum resource.

This distinction is also important for the study of prejudice. Criticism of a government is a normal part of political debate. Collective blame aimed at an ethnic or religious population is something different. The historical lesson of antisemitism is that political grievances can become dangerous when they are transferred from institutions or leaders to an entire people. A parallel principle applies to Palestinians and Muslims: political disagreement should not become a justification for treating a population as collectively suspect or less deserving of safety.

The value of including Zionism in this paper is therefore not to resolve the Israeli-Palestinian conflict. Such a goal would be far beyond the scope of one survivor interview. Instead, Karin's perspective demonstrates how historical memory enters contemporary political identity. The Holocaust is not merely a past event remembered by Jewish communities; for many Jews it shapes ideas about vulnerability, sovereignty, and the consequences of relying on others for protection. Palestinian historical memory likewise shapes ideas about dispossession, homeland, and self-determination. Understanding the power of those memories does not eliminate disagreement, but it can make disagreement more historically informed and less dehumanizing.

The Limits and Responsibilities of Historical Comparison

The themes that emerge from Karin's testimony invite comparison with contemporary issues, but they also demonstrate why comparison must be restrained. Holocaust memory is sometimes used rhetorically to describe almost any disliked political development. That practice can flatten the specific historical conditions of Nazi genocide and replace analysis with moral labeling. A more responsible approach asks not whether two events are "the same," but whether a particular mechanism—such as legal exclusion, collective blame, propaganda, or social dehumanization—appears in both contexts and, if so, how the contexts differ.

Karin's case suggests several mechanisms worth watching without claiming inevitable repetition. The first is the conversion of stereotypes into law. The second is the use of group identity to override individual behavior. The third is the normalization of unequal treatment

through gradual steps. The fourth is the social withdrawal of people who once regarded members of a minority as neighbors. The fifth is the dependence of vulnerable people on the willingness of others to intervene. None of these mechanisms belongs exclusively to one time or place, but their presence alone does not prove that a society is moving toward genocide.

This distinction protects both historical accuracy and moral seriousness. If every controversy is described as another Holocaust, the term loses analytical value. If the Holocaust is treated as so unique that no aspect of it can inform later societies, historical study loses much of its preventative purpose. The challenge is to hold both truths at once: the Holocaust was a historically specific genocide, and the processes of prejudice and exclusion that contributed to it remain human possibilities.

The same principle applies to the personal comparison between Karin's history and my grandparents' immigration. Their experiences should not be merged into one story of hardship. Yet placing them in conversation can illuminate how differently people experience the relationship between home and political belonging. Karin learned that a birthplace could stop functioning as a home when the state defined her as an enemy. My grandparents left one home and attempted to construct another. The comparison is meaningful precisely because the differences remain visible.

Conclusion

Gabriella Karin's testimony demonstrates the value of studying the Holocaust through both historical documentation and individual memory. The documentary record establishes the political transformation of wartime Slovakia: antisemitic legislation, the Jewish Code, deportation, and mass murder. Karin's memories show how that transformation was experienced inside one family. A child who had once understood Bratislava as home came to depend on false papers, secrecy, religious institutions willing to conceal her, and the courage of a lawyer willing to hide Jews in his apartment.

Her story supports the central argument of this paper: extreme persecution is made more possible when exclusion becomes ordinary. The Holocaust did not begin when the first deportation train departed Slovakia. It was preceded by a social and legal process that separated Jews from their neighbors, narrowed their rights, removed them from public life, and taught society to accept their unequal status. Karin's experience also demonstrates that assimilation cannot guarantee safety when belonging is defined through ancestry. Her family could participate in Slovak life and still be reclassified as outsiders.

At the same time, Karin's survival reveals that individuals retain moral agency even within oppressive systems. The people who protected her could not control the government, but they could decide whether its categories determined their own behavior. Their choices are part of the history of the Holocaust just as surely as the laws and deportations are. Remembering them does not soften the history of genocide; it clarifies what resistance can look like at the level of ordinary life.

The continued importance of survivor testimony follows from this connection between history and choice. Oral testimony cannot replace archival evidence, and memory must be

evaluated with the same care historians bring to other primary sources. Yet testimony makes visible what institutional records often cannot: fear, uncertainty, the feeling of losing home, and the meaning attached to another person's decision to help. As survivors age, preserving those perspectives becomes an increasingly important part of Holocaust scholarship and education.

Finally, Karin's testimony changed how I understand contemporary questions of antisemitism, immigration, Zionism, and belonging. It did not provide simple political answers. Instead, it made simple answers harder to accept. Jewish historical trauma helps explain why security and self-determination carry profound meaning for many Jews. Palestinian displacement must also be recognized as a central historical reality. Immigration debates can address law and policy without reducing immigrants to stereotypes. Opposition to antisemitism can be strongest when it is connected to a broader rejection of dehumanization.

"Never Again" is meaningful only if it is more than a slogan attached to the past. It should not be used to declare every modern conflict another Holocaust, but neither should it be reduced to a ceremonial phrase. Karin's life suggests a more demanding interpretation: remember how exclusion became normal, remember what happened when citizenship ceased to protect a minority, remember the people who chose to help, and recognize that historical memory creates obligations for those who inherit it. The point of testimony is not only to know what happened. It is to become more capable of recognizing the choices that shape what happens next.

Works Cited

Holocaust Museum LA. "Survivor Talk with Gabriella Karin." Holocaust Museum LA, 28 Dec. 2023, www.holocaustmuseumla.org/event-details/survivor-talk-with-gabriella-karin.

Karin, Gabriella. Zoom interview with the author. July 2026.

Lipsitz, Neal. "Oral Memoirs: The Testimony of Holocaust Survivors." Oxford Research Encyclopedia of Psychology, 27 Aug. 2020, doi.org/10.1093/acrefore/9780190236557.013.480.

Museum of Tolerance. "Gabriella Karin." Museum of Tolerance, 7 June 2026, museumoftolerance.com/event/wtt-speaker-gabriella-karin.

Museum of Tolerance. "Witness to Truth." Museum of Tolerance, museumoftolerance.com/visit/exhibits/witness-to-truth.

Pew Research Center. "How Much Discrimination Do Americans Say Groups Face in the U.S.?" 20 May 2025, www.pewresearch.org/politics/2025/05/20/how-much-discrimination-do-americans-say-groups-face-in-the-u-s/.

Passel, Jeffrey S., D'Vera Cohn, and Ana Gonzalez-Barrera. "II. Migration Between the U.S. and Mexico." Pew Research Center, 23 Apr. 2012, www.pewresearch.org/hispanic/2012/04/23/ii-migration-between-the-u-s-and-mexico/.

Pew Research Center. "Chapter 1: The Nation's Immigration Laws, 1920 to Today." Modern Immigration Wave Brings 59 Million to U.S., 28 Sept. 2015, www.pewresearch.org/race-and-ethnicity/2015/09/28/chapter-1-the-nations-immigration-laws-1920-to-today/.

United Nations Information System on the Question of Palestine. "Palestinian Refugee Factbox/Timeline – IRIN News Article." 22 June 2010, www.un.org/unispal/document/auto-insert-199544/.

United States Holocaust Memorial Museum. "Axis Powers and the Holocaust." Holocaust Encyclopedia, encyclopedia.ushmm.org/content/en/article/axis-powers-and-the-holocaust.

United States Holocaust Memorial Museum. "Historian Peter Black Describes the Value of Eyewitness Testimony." Holocaust Encyclopedia, encyclopedia.ushmm.org/content/en/oral-history/historian-peter-black-describes-the-value-of-eyewitness-testimony.

United States Holocaust Memorial Museum. "Holocaust Survivors and the Establishment of the State of Israel." Holocaust Encyclopedia, encyclopedia.ushmm.org/content/en/article/postwar-refugee-crisis-and-the-establishment-of-the-state-of-israel.

Yad Vashem. "Deportations." Bratislava During the Holocaust, wwv.yadvashem.org/yv/en/exhibitions/communities/bratislava/deportations.asp.

Yad Vashem. "Righteous Among the Nations Honored by Yad Vashem: Slovakia." Yad Vashem, wwv.yadvashem.org/yv/pdf-drupal/slovakia.pdf.

Procedural and Conceptual Knowledge: a Chronological Relation in Basic Derivatives
By Quinn Harman

Introduction

Education is a constantly changing and evolving topic. Whether it's the topic being taught at different grade levels or simply different styles of teaching across countries, there is always a common thread: conceptual and procedural knowledge. Conceptual knowledge can be defined as "the knowledge of concepts, which are abstract and general principles," and procedural knowledge can be defined as "the knowledge of procedures" (Rittle-Johnson et al.). These two types of knowledge appear everywhere within education, and are generally agreed to be closely linked together in a multitude of ways. The linkage between the two types of knowledge has been studied in depth in many areas, namely STEM fields, but primarily mathematics. The effects that these types of knowledge have on education have been studied significantly, such as how students with a better conceptual understanding of a mathematical topic had improved procedural understanding in the long term compared to students with a lower conceptual understanding (Rittle-Johnson & Alibali).

Literature Review

To fully understand the effects that conceptual and procedural knowledge have on education, however, the linkage between the two must be understood. To start, does one type of knowledge lead to another? From Rittle-Johnson and Alibali, teaching one type of knowledge to students, in their study using specifically 4th and 5th grade participants, led to a greater understanding of the other type of knowledge. While this study hints at a linkage between the two types of knowledge in conceptual supporting procedural knowledge and vice versa, the study specifically states that it may not be equal both ways. Similarly, another study from Silver agrees with Rittle-Johnson and Alibali that while linkage between the two types of knowledge certainly exists, it may not always be equal (Silver). More specifically, it seems that procedural knowledge does not support conceptual knowledge in the same way as conceptual knowledge supports procedural knowledge. However, Rittle-Johnson, Schneider, and Star believe that the two types of knowledge support each other bidirectionally (both sides support each other equally) through significant testing and analysis from past studies (Rittle-Johnson et al.). This contrasts significantly from other beliefs, stating that one side supports the other more. However, both sides hold valid points as there is significant research across many sources suggesting that it is more one-sided than bidirectional, such as past studies even from Rittle-Johnson and other colleagues. Still, the general baseline for all of them, regardless of their stance, is that conceptual knowledge does support procedural knowledge, and there are deep linkages between the two types of knowledge.

Both conceptual and procedural knowledge support each other, either bidirectionally or omnidirectionally, but it's important to understand how. Multiple studies, such as from Rittle-Johnson, Bethany & Siegler, Robert & Alibali, and Martha, discuss the idea that the two

types of knowledge develop on each other iteratively. What this means, as defined by Rittle-Johnson and colleagues, is that when one type of knowledge is taught, it improves the other type, which in turn leads to improvements in the previous type (Rittle-Johnson et al.). This completes a type of iterative cycle, where the two types of knowledge are always improving each other to an extent. This is important to understand as it helps to explain the linkages slightly more, but also raises the question of "does one side of this process put more work in than the other?" While very little research has been done on this, it can be inferred from studies such as Rittle-Johnson and Alibali that conceptual knowledge may have more of an effect on the process than procedural knowledge does.

There are also specific areas within education and in general where conceptual and procedural knowledge have increased or altered effects, namely in STEM fields. In STEM fields, conceptual and procedural knowledge are integral in a multitude of ways, especially in mathematics. Firstly, conceptual and procedural knowledge play an important role in mathematics due to the complex topics where both a procedural and conceptual understanding are required in order to fully understand a topic, as shown by Euler in a study in Singapore testing the linkage of knowledge in STEM education (Euler, 105-117). These two types of understanding intertwine closely of whether, such as a geometric proof, when a conceptual understanding is required to read and create the proof, and a procedural understanding is required to follow through and write/create the proof (Schoenfeld, 225-264). The two types of knowledge have a significant linkage between them, which has led researchers to ask the question if they are actually separable from each other. This question is vitally important due to the fact that if they were to be found inseparable, then many theories of intertwining between the two types of knowledge would be affected significantly. The study from Lenz, Dreher, Holzäpfel, and Wittmann focused on whether the two types of knowledge can be separated when dealing with the education of fractions, due to the previous consensus of past research and studies that showed significant linkages between the types of knowledge in mathematics. Through the study, the authors concluded that while the two types of knowledge showed significant linkage and similarity, they were empirically separable when dealing with fractions (Lenz et al.). However, this raises the question of whether the two types of knowledge are still separable when dealing with other and more complex topics? Research from many other authors suggests that conceptual and procedural knowledge may be too closely tied together to be separated when dealing with complex topics. This conjoins with the thought of how both types of knowledge support each other, as suggested by previous studies, as if the two types of knowledge were not empirically separable, then the effects might not be concrete. Another example of an effect on conceptual and procedural knowledge in a specific field can be seen in a study by Beeler and colleagues, where they conducted a study to see the effects conceptual and procedural knowledge had in the medical field with visual learning (Beeler et al.). Their results combined similarly with Rittle-Johnson, Schneider, & Star as the studies supported the idea that procedural knowledge and conceptual knowledge were both supporting the students' learning, albeit in different ways.

It’s also suggested that there can be significant impacts on education depending on the teacher's knowledge and understanding of both aspects of knowledge in a given topic and the separation between the two types of knowledge. Firstly, from Hurrell in “Conceptual knowledge OR Procedural knowledge OR Conceptual knowledge AND Procedural knowledge: Why the conjunction is important for teachers”, the author states that the difference between the and/or between conceptual and procedural knowledge is vitally important to teachers (Hurrell). This is because, similarly to previous claims by other authors such as Rittle-Johnson, Schneider, & Star, the two types of knowledge show significant linkage, and when teaching both types of knowledge, students' understanding improves overall.

Over the past decade, there has been a noticeable decline in students’ understanding of topics, namely in mathematics, in the United States. The NAEP (National Assessment of Education Progress) scores have declined noticeably over the last decade, specifically in mathematics. Taking the scores from the assessments (as shown in the diagram) from 2013 to 2024, there is a significant drop in average national score, from 153 down to 147 (scores out of 300). The NAEP classifies this as a significant drop from year to year, showing a decrease in understanding from students (National Assessment of Education Progress). Another reason for these significant drops in student understanding is the COVID-19 pandemic. From Zhdanov, Sergei and colleagues, school closure due to COVID-19 did cause significant learning loss among students during that time, which, as shown by other studies such as the report from the NAEP, has not been recovered from (Zhdanov et al.). If new strategies are not implemented, this learning loss could simply continue to increase over time, leading to significant issues for future generations.

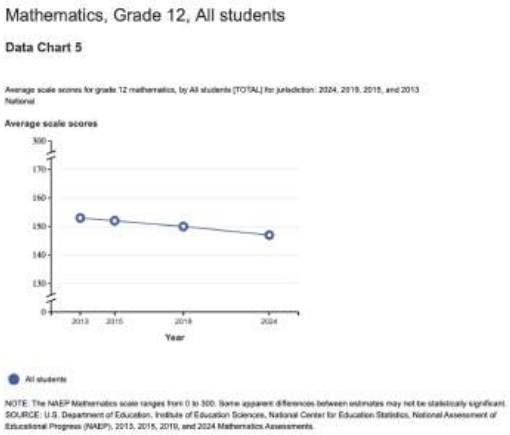


Fig 1: Graph of average NAEP Scores for grade 12 mathematics students from 2013-2014

Despite the clear conjunction between conceptual and procedural knowledge and its importance, there is still very little information about the effects that the order of teaching these two types of knowledge brings. Many researchers are in agreement about the consistent effects of these two types of knowledge when combined, and significant studies have been done on it. However, these researchers and papers also acknowledge the fact that while the linkages are

clear, it is unclear what effect the order of teaching would have. Logically, with the specific linkages of a large portion of the research suggesting that conceptual knowledge leads to procedural knowledge, it would be beneficial to teach the conceptual portion of a topic before the procedural aspect.

Methodology

In order to conclude whether conceptual or procedural knowledge should be taught first, I conducted an experiment using students enrolled in AP Precalculus that tested procedural and conceptual knowledge order with basic derivative rules. An experiment is preferred to answer the research question as previous studies and experiments have been unable to provide substantial evidence to support any answers, and therefore, new research must be conducted. Because students under 18 (sophomores and juniors) are the participants in the experiment, consent forms were collected before the experiment could be completed, and no student names were collected during data collection, so pure confidentiality is maintained throughout the experiment, and no names can be placed on scores. Students in AP Precalculus are preferred for this experiment, as the average student in AP Precalculus will have no knowledge over derivative rules, but also will still have the foundation needed in order to comprehend the topic.

The experiment is divided into 2 separate groups, with 2 stages each. For each group, they will be shown two identical videos, one over conceptual knowledge for basic derivative rules, and one on procedural knowledge. Group 1 will be shown the video covering procedural knowledge, and after a 1-minute break, will be shown the video covering conceptual knowledge. Group 2 will first be shown the video covering conceptual knowledge, and after an identical 1 minute break, be shown the video covering procedural knowledge. After being shown both videos, each student in both groups will be given an identical test over the content covered in both videos. Tests are then collected individually, and scores are averaged and recorded. The study consists of 6 students in total, 3 in each group. This is concurrent with similar studies from Gün Şahin, Zühal & Gürbüz, Ramazan, where they used 10 students in total to cover the focus of their study, the use of short films in procedural and conceptual knowledge teaching (Zühal & Ramazan). This study uses 3 per group to approximate an average score, cover both areas of the study, and stay concurrent with previous studies.

The test itself and videos cover basic derivative rules. This includes the power rule, defined as: $\frac{d}{dx}x^n = nx^{n-1}$, the constant rule: $\frac{d}{dx}(c) = 0$, the constant multiple rule: $\frac{d}{dx}[cf(x)] = [cf'(x)]$, and the sum and difference rules: $\frac{d}{dx}[f(x) \pm g(x)] = f'(x) \pm g'(x)$ developed by Issaac Newton, Johann Bernoulli and Gottfried Leibniz. The conceptual video explains these rules without actually working through them, along with the limit definition of a derivative and what instantaneous rate of change through using a diagram of a graph with two points converging. The procedural video simply contains 10 examples, increasing in difficulty while including different rules, and manually working through each problem with no commentary, but simply solving the problems through the rules that are used in the conceptual video without any further explanation.

The experiment was designed through specific videos in order to be able to replicate the experiment as identically as possible by using the exact same videos. While this experiment focuses on students in AP Precalculus in the Kansas City area, these videos allow for AP Precalculus students in any area to be able to participate in the same experiment if replicated. The design also allows for new videos to be made at lower or higher grade levels, and see if the results differ based on grade level.

Analysis of results

To accurately analyze the data that was gathered through the experiment, we must look at it through two angles, both quantitatively to observe a numerical correlation and qualitatively to judge an accurate level of learning. When looking at the data collected from the experiment, it defies the original hypothesis, as the participants who were shown the procedural video first, on average, performed better than the participants shown the conceptual video first. The average score for participants who were shown the procedural video first was a 3.33/10 or a 33.3%, and the average score for students shown the conceptual video first was a 5/10 or 50% (Shown below in Figure 2).

Conceptual Shown First			Procedural Shown First		
Student #	Test Scores	Average Score	Student #	Test Scores	Average Score
Student 1	3 / 10		Student 1	9 / 10	
Student 2	4 / 10		Student 2	1 / 10	
Student 3	8 / 10		Student 3	0 / 10	
		5/10 (50%)			3.3/10 (33%)

Fig 2: List of test scores across experiment, including averages

While these scores may seem extremely low compared to normal American grading standards (0-59%=F, 60-69%=D, 70-79%=C, 80-89%=B, and 90-100%=A), this is somewhat to be expected. This is because while the participants all came from AP Precalculus have the basis of knowledge needed in comparison to taking multiple weeks in an average high school curriculum, the topic was introduced in a very short period of time. The difference in average score also seems small and possibly inconsequential due to the small sample size. However, the trend shown in the data still provides strong empirical evidence for an advantage to teaching conceptual and procedural knowledge in a specific chronological order. To do this more specifically, we can analyze which rules were the most well understood between the two groups. For the power rule, group 1 (procedural video first, then conceptual) completed the problems that were just the power rule with 80% accuracy, while group 2 (conceptual video first, then procedural) had a 66% accuracy. For the addition and subtraction rules, both groups completed the basic procedure of separating the terms with 100% accuracy. Next, for the constant multiple rule, group 1 completed problems including it at a 75% accuracy, and group 2 completed the problems with a 66% accuracy. You can see that throughout this, while margins may be low, group 1 consistently completed rules with a higher accuracy than group 2. Even though the quantitative results and margins between accuracy percentages between the groups are extremely

low (mean scores being low to be expected as the experiment covers a higher-level topic), the qualitative results still show a significant advantage to one method of learning, and that was through learning conceptual before procedural, following the original hypothesis. The biggest example of this is the constant rule, while very easily completed procedurally, it can be confusing conceptually when also trying to understand multiple other new rules at the same time. On the test, there were two problems that were solely the constant rule and included no other rules, and for the batch of tests that observed the procedural video before the conceptual only correctly answered these questions 75% of the time, while the reverse answered them correctly 100% of the time. While this could be a difference simply due to a small sample size in the experiment, this is not the only example of this phenomenon within the results. When analyzing the results of the batch of tests that watched the conceptual video first, on average, their answers were closer to being correct than the other group, mostly being calculation errors—such as multiplying by 1/7 instead of by 7 and errors based off negative exponents when translating the power rule and other rules in subsequent problems—while the other batch showed mostly conceptual errors with many blank answers and wrong processes instead of calculation errors. The final way that data and trends can be extrapolated from this study is through analyzing which rules were applied falsely throughout both groups. The most common example of this is the concept of the product rule being used on problems containing the constant multiple rule and the product rule (likely originating from confusion on what to do with the exponent when utilizing the constant multiple rule). This also occurred with the power rule with negative exponents, and this can be attributed to both a general misunderstanding of the power rule when concerning negative exponents and a mistake due to not spending a ton of time on a given problem and not converting a term such as $\frac{7}{x^4}$ to $7x^{-4}$ and simply subtracting from 4 instead of understanding that it is, in reality, negative 4. Overall, all these methods of analyzing the results of the study ultimately support the trend shown by looking at solely the mean scores.

Conclusion/Findings

First, these results reinforce the original hypothesis that using a chronological order to teach the conceptual knowledge first leads to a higher level of understanding, at least for the topic of basic derivative rules. Both the quantitative trend of scores and the qualitative analysis of consistent and common mistakes contribute to the reinforcement of the original hypothesis. The strength of this reinforcement is fairly strong, especially when analyzing the qualitative results. However, the quantitative results don't immediately show as strong of a trend, but they do still support the original hypothesis by showing an advantage to teaching conceptual knowledge first. These methods of analyzing both contribute to a new level of understanding that had previously not been explored within the study of procedural and conceptual knowledge. Furthermore, it also supports past ideas of studies by Rittle-Johnson and other colleagues, suggesting that conceptual knowledge may have more of an impact on procedural knowledge than the inverse scenario. However, it does not necessarily support the studies that say procedural knowledge has little to no effect on conceptual knowledge due to the fact that this

experiment does not include a way to test students' understanding apart from separating conceptual and procedural knowledge. Further research would have to be done for similar hypotheses to be further supported or proven. Finally, the trend shown in the study contributes to the idea that conceptual and procedural knowledge are empirically separable (shown simply by the fact of there were different results between the two groups), which could aid in future research concerning conceptual and procedural knowledge, as the idea of whether the two types of knowledge were empirically separable was still under question.

Limitations

Despite primarily clear results from the majority of the study, some limitations must be addressed to understand the context of the results. First, the trend may change with larger sample sizes. In the context of larger research, the sample size of this study is quite small due to the specificity of requirements for participants. While the qualitative analysis of data suggests that the trend of a chronological order having a significant effect will stay consistent, it could end up showing something different with a larger sample size. Secondly, this study only observes a very specific topic by addressing basic derivative rules. Many different math courses (or even science courses) may see different results when replicating the experiment with a different focus, and the study cannot be generalized to all math topics, nor can it be generalized to any science or physics topics. Finally, the results of this study cannot be generalized to all students either. Similarly to the topic, the study pulls a very specific group (AP Precalculus students at a public high school in the Midwestern region of the United State), and a replicated version of the experiment may show different results when analyzing college students or even just students in other states or countries. Thirdly, the age of the students could also have an impact on the results of future studies, as at different levels of education (in the United States, primarily elementary, middle, and high school as well as college), often use very different methods of teaching, which could lead to a difference in the trend shown by data from a similar study. While a change in these variables has a chance of causing a notable change in the results of the study, it is mostly unclear on what exactly it would do, especially when changing the topic being studied. Finally, the trend shown by the study is mostly applicable to other topics in math education when aspects of the topic are empirically separable (such as fractions from a study by Lenz, Dreher, Holzäpfel, and Wittmann where they found a similar result to the study in this paper, concluding that the types of knowledge were separable through an experiment) but may be difficult when topics are not as easily separable as the study will not be as accurate to the original and may produce a different trend. However, this new trend may also support other hypotheses, which is why it is extremely important that more research on this topic must be completed.

Further Research and Implications

This research shows a possibly extremely important trend that could lead to significant improvement in the education system, however, there is still work that must be done. To introduce these concepts into curriculums on a larger scale, similar experiments would need to be

replicated on different topics and over longer periods of time. Once the results can be proved and replicated on a larger scale, it could lead to creating a significant advantage for many educators and their students in learning math concepts. This is especially important due to the fact that the average math understanding in the United States is and has been declining ever since COVID-19 first broke out. A study from Vegas, Patrinos, and Carter-Rau, in 2022, analyzed national test scores and determined that understanding was significantly declining. They strongly recommended that solutions must be found to help combat this learning loss among students (Vegas et al.). Knowing this information, any advantages educators can use to help their students understand more to combat the ever continuing decrease in students' understanding is an advantage that should be researched and investigated to its fullest. However, to fully apply this phenomenon and ideas to the American education system, further studies must be completed, such as testing different topics and age groups, other areas that differ in median income, and all other variables that can be changed within the study. This reinforces the implication that this study already begins, and could bring significant change and help to a declining education system that has received many complaints and concerns from students, parents, and teachers alike. Furthermore, the results of this study could hold significant advantages for students with mathematical disabilities. In a study by Lee in 2024, he analyzed longitudinal studies and found that, on a general basis, students without mathematical disabilities benefited from conceptual instruction, while students with mathematical disabilities benefited more from procedural instruction (Lee). Chronological order may also have an effect and is another reason that this phenomenon could be helpful in education for many students. However, it is uniquely useful when being used in STEM (science, technology, engineering, and math) classes, as not only does the idea of separating procedure and concepts become significantly easier, but STEM is rapidly advancing. From Euler, as previously mentioned, conceptual and procedural knowledge are integral within STEM education, especially at higher levels of education (Euler). The results of this study could be used frequently within those high levels of STEM classes, such as math classes, including derivatives as included in the study, and could benefit in a large way. Along with increasing education in STEM, as the future approaches and new technologies are made every single day, it becomes increasingly important to teach the new generation in STEM topics. Even without the current decline in mathematics understanding within the current generation of students, the idea of procedural and conceptual knowledge is somewhat unique to STEM classes and curriculum as many other areas struggle to have quite as easily defined differences between the two types of knowledge (such as english courses having a similar thing by teaching grammar and then having students write essays, however, this does not translate quite as easily as math or science rules into solving problems). Students must be taught in the most efficient ways possible as society continues to progress and become more fast-paced, and showing a linkage between conceptual and procedural knowledge and knowing that there is a significant effect of chronological order will only help and advance STEM education for the future of our society.

Works Cited

Beeler, N., Ziegler, E., Volz, A., Navarini, A. A., & Kapur, M. "The effects of procedural and conceptual knowledge on visual learning." Advances in Health Sciences Education, 29(4), 1243–1263. https://, 2023. https://doi.org/.org/10.1007/s10459-023-10304-0

Euler, M. "The Mechanics of Creative Cognition: Orchestrating the Productive Interplay of Procedural and Conceptual Knowledge in STEM Education." Proceedings of the Singapore National Academy of Science, 15(02), 105–117. https://, 2021. https://doi.org/.org/10.1142/s259172262140010x

GUN SAHIN, Z., & GURBUZ, R. "The effect of supporting realistic mathematics education with short films on conceptual and procedural knowledge." Acta Didactica Napocensia, 15(2), 83–110. https://, 2022. https://doi.org/.org/10.24193/adn.15.2.6

Harry Anthony Patrinos, Vegas, E., & Rohan Carter-Rau. "An Analysis of COVID-19 Student Learning Loss." Oxford Research Encyclopedia of Economics and Finance. https://, 2023. https://doi.org/.org/10.1093/acrefore/9780190625979.013.893

Hurrell, D. "Conceptual Knowledge OR Procedural Knowledge or Conceptual Knowledge AND Procedural Knowledge: Why the Conjunction is Important to Teachers." Australian Journal of Teacher Education, 46(2), 57–71. https://, 2021. https://doi.org/.org/10.14221/ajte.2021v46n2.4

Lee, J. "Concept-Focused and Procedure-Focused Instruction on the Algebra Performance of Grade 9 Students With and Without Mathematics Difficulty." Journal of Learning Disabilities. https://, 2024. https://doi.org/.org/10.1177/00222194241249960

Lenz, K., Dreher, A., Holzäpfel, L., & Wittmann, G. "Are conceptual knowledge and procedural knowledge empirically separable?." The case of fractions. British Journal of Educational Psychology, 90(3). https://, 2019. https://doi.org/.org/10.1111/bjep.12333

National Assessment of Education Progress. "Mathematics, Grade 12, All students https://www.nationsreportcard.gov/ndecore/xplore/NDE." , 2026.

Rittle-Johnson, B., & Alibali, M. W. "Conceptual and procedural knowledge of mathematics: Does one lead to the other?." Journal of Educational Psychology, 91(1), 175–189. https://, 1999. https://doi.org/.org/10.1037/0022-0663.91.1.175

Rittle-Johnson, B., Schneider, M., & Star, J. R. "Not a One-Way Street: Bidirectional Relations Between Procedural and Conceptual Knowledge of Mathematics." Educational Psychology Review, 27(4), 587–597. https://, 2015. https://doi.org/.org/10.1007/s10648-015-9302-x

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. "Developing conceptual understanding and procedural skill in mathematics: An iterative process." Journal of Educational Psychology, 93(2), 346–362. https://, 2001. https://doi.org/.org/10.1037//0022-0663.93.2.346

Schoenfeld, A. "On having and using geometric knowledge." In Conceptual and Procedural Knowledge: The Case of Mathematics. Routledge., 1986.

Silver, E. "Using conceptual and procedural knowledge: A focus on relationships." In Conceptual and Procedural Knowledge: The Case of Mathematics. Routledge., 1986.
Zhdanov, S. P., Baranova, K. M., Udina, N., Terpugov, A. E., Lobanova, E. V., & Zakharova, O. V. "Analysis of Learning Losses of Students During the COVID-19 Pandemic." Contemporary Educational Technology, 14(3), ep369. https://, 2022. https://doi.org/.org/10.30935/cedtech/11812

The Evidence of the Dead: Material Truth and Political Memory in Lucan's *Pharsalia*
By Derek Hu

Abstract

Lucan's *Pharsalia* presents civil war not only as a political and military crisis, but also as a crisis of knowledge. Because both sides of the conflict are Roman, conventional signs of political identity and authority become increasingly unreliable, creating a need for alternative forms of certainty. This paper argues that Lucan repeatedly turns the dead body into a form of evidence. Through the mutilated victims of Sulla, the resurrected corpse of Ericto, and the preserved head of Pompey, bodies are used to establish identity, provide testimony, and confirm political facts. However, the same bodies that appear to offer material certainty are also mutilated, manipulated, preserved, and interpreted by the living, which undermines their apparent reliability. Lucan therefore develops a paradox in which civil war increasingly depends upon physical evidence at the same time that the violence destroys the stability of that evidence. The poem ultimately shifts the problem of preserving truth from bodies to graves, inscriptions, landscapes, and finally poetry itself. By following this progression, this paper argues that *Pharsalia* explores not only how civil war destroys Roman lives and political institutions, but also how historical truth is constructed, preserved, and contested after violence has ended.

Introduction

In Lucan's *Pharsalia*, the dead rarely disappear once violence has ended. Bodies continue to occupy the poem long after death has occurred, as they are mutilated, searched, resurrected, displayed, preserved, buried, named, and remembered. Lucan often devotes more attention to what happens to a victim after death rather than the act that actually led to their death, making the corpse an enduring presence rather than a consequence of battle. This treatment is especially notable in a work concerned with a conflict whose political outcome was known long before Lucan began writing. Caesar had won, the Republic had fallen, and the political system that followed was the same one Lucan was living in. Yet *Pharsalia* returns to the true aftermath that the victorious history left behind, giving the reader a different view of Rome from the one that's seen in Virgil's more commonly known work, Aeneid, which presents the costs of expansion and military success as an inevitable part of Roman destiny Rather than using warfare to trace the growth and destiny of Rome, Lucan instead analyzes how Rome becomes both victor and victim, where military achievement produces corpses rather than national progress. In this manner, Lucan transforms the traditional possibilities of epic poetry. Instead of allowing the civil conflict to pass into the past, Lucan repeatedly recovers its severed bodies, neglected remains, unfinished burials, and ruined landscapes. The persistence of the dead therefore raises a question beyond why Lucan depicts civil war so violently: why must its victims continue to be handled, preserved, and brought before the reader even after their deaths? This paper argues that the dead bodies transform from reminders of violence into unreliable sources of information. When civil war destroys the institutions that establish truth and identity, the people turn to the corpses to

identify the once-living, explain what had happened, prove the success or failure of political decisions, and to preserve the memories of their past. However, Lucan shows that the bodies are not reliable sources of information, as they can be mutilated or resurrected and subject to manipulation, but that through poetry, he is ultimately able to preserve what the bodies cannot.

This concern with what remains after destruction is closely connected to Lucan as a writer. Physical traces of the past are temporary. Bodies decay, graves collapse, monuments disappear and even famous locations become impossible to recognize. Lucan presents this problem when Caesar visits the ruins of Troy. The city's physical remains have deteriorated that Caesar unknowingly walks across places once associated with figures like Hector and requires a guide to identify what remains around him. Lucan responds by praising the "sacred labour of the poet," (IX.980), who can rescue people and events from being forgotten in time. Writing consequently becomes a tool to determine what aspects of specific events remain available to later generations and prevents political victory from possessing complete control over historical memory. Instead of placing poetry entirely for heroic commemoration, Lucan uses his role as a poet to preserve what heroic narratives might prefer to look over: defeat, mutilation, guilt, and the victims left behind by Roman success. Lucan therefore openly contests what the events that are fundamental to Rome should mean. Writing during the reign of Nero, Lucan returns attention not to those who prevailed but to the human remains and unresolved violence produced during the process. His later involvement in the Pisonian conspiracy and forced suicide under Nero gives additional importance to a work so concerned with advocating for what can survive the consolidation of political power. In this way, Lucan retains a form of authority as he maintains the ability to determine what later readers are forced to remember about that victory.

This ability to preserve competing versions of the past also explains why the poem remains relevant beyond the political circumstances of first-century Rome. Modern societies continue to construct public memories of wars, political violence, and national victories through monuments, official accounts, photographs, testimony, and literature. The problem Lucan presents is therefore generalized throughout the world as the group that wins a conflict often possesses the power to explain the conflict afterward, while the physical and written remains of the conflict may still provide details that are disregarded or simplified. The poem remains significant because it demonstrates that control over history is not identical to control over memory. Material remains and writing can continue to challenge an established account long after the political struggle itself has ended.

Scholarship has done extensive work on the language and purpose of Lucan's writing with respect to death and corpses. Dinter studies the use of body imagery within the epic and argues that corporeal imagery helps fuse the fragmented sections together (Dinter 50). McClellan studies interactions with the dead—corpse abuse, denial of burial, improper funerals, and corpses—to analyze the disturbed relationship between life and death arising from the conflict (167). Mebane analyzes the decapitation of Pompey specifically as a representation of Rome's political state, describing the implications of separating the "head" of Rome from its "body" (191-196). Rather than analyzing what Lucan's corpses represent, this essay explores how the

living use the dead bodies to make sense of the destruction of war. Through identifying the dead, extracting information from the bodies, using them as proof, and preserving them, Lucan uses the corpses as an object through which he explores the war-torn past of the Roman people.

Civil War and The Problem of Truth

As opposed to a standard conflict between enemy and enemy, civil war poses a unique epistemological problem: Romans can no longer rely upon conventional political distinctions and authorities to determine reality. At the major battle on the plains of Pharsalus, both sides are fundamentally composed of Romans; Lucan repeatedly emphasizes that fathers and brothers face one another across the battlefield. This establishes the central problem since kinship, like political affiliation, offers no reliable way to tell an ally from the enemy on sight. If the war is defined by the absence of a consistent "enemy", then the terms that normally organize conflict, and by extension truth, no longer apply. In turn, some other means of establishing certainty becomes necessary.

In an ordinary Roman conflict, standards and eagles announce which side is Rome, signifying to friend and foe alike the glory and authority of the Roman state. These signs work as distinctive signs purely because they are exclusive to only one army. As the people of Marseilles tell Caesar, they will take up arms against any foreign enemy, but beg him to "leave the dread eagles and menacing standards far from our city" the moment the conflict turns civil (III.309-310). Before the grand battle on the plains of Pharsalus, Lucan describes the standards as having "wept real tears, for until Pharsalia, they had signified Rome and the State."(VII.162-163), foreshadowing the moment where these objects will cease to mean anything. The eagle's presence usually declares that Rome is here, but when both armies carry Roman eagles, the sign physically persists while its meaning collapses into contradiction. Because it now marks both armies, it therefore marks neither. Such an iconic symbol relies on its uniqueness and scarcity; duplicated, it has no purpose. However, the physical body appears to contain no such vulnerability. A corpse cannot be doubled or shared between opposing claims the way a standard claim, allowing it to retain an evidentiary power that political signs lose in similar situations. When political language can no longer reliably distinguish Rome from Rome's enemy, the body appears to promise something more immediate: recognition.

Recognition

Lucan first establishes the evidentiary body through what might be called the paradox of recognizable mutilation: The more that the victim is mutilated, the harder it is for the victim to be identified; however, political violence is only effective when the victim can still be identified as the object of punishment. During Lucan's recollection of Sulla's tyrannous reign, violence is widely encouraged and sanctioned, resulting in widespread chaos, such as "[a] servant drove the accursed blade through his master's body; sons were drenched in their fathers' blood and fought for the privilege of beheading them; brother killed brother for a price."(II. 148-151). This violence turned kin on kin, not only through quick killings, but through pain and torture that

would often cause bodies to become extremely mutilated and disfigured. An old man reflects upon his personal experiences as "I myself remember how I searched the corpses murdered in Sulla's 'peace', trying to recognise the shapeless features of my murdered brother, to rescue him from the fire. Round all the decapitated dead I went, seeking a neck to fit that severed head."(II. 167-172). Here, the corpse is forced to perform a function that is normally the job of the living, which is confirming its own identity. Without a voice, social position, the ability to self name, the only thing that remains is physical recognition. When this factor is heavily altered, it becomes nearly impossible to identify a loved one, forcing Lucan to try and fit a head to a neck. In this sense, the physical body serves not only as a reminder of the savagery of Sulla, but a means of identification.

Lucan also extends this to politicians, who are not exempt from this treatment, and can often be subject to far worse disfigurement. He describes the mutilation of Gratidianus, a fierce supporter of Sulla's rival Gaius Marius, by Sulla's supporters: "His arms were wrenched from the shoulders, his tongue, severed, quivered and beat the empty air with mute motion; one man pared the ears, another the nostrils of the curved nose; a third man dragged the eyeballs from their sockets, and when the eyes had witnessed the limbs' fate, cut them both free."(II. 178-182). As a political opponent, it should be expected that Sulla would want to keep him recognizable as proof of his death. Yet, this did not happen, rendering the political killing partially ineffective. Under a regime such as Sulla's, murder is not just about killing somebody, but making an example out of them and doing so in such an extreme and often savage manner. Therefore, Lucan reveals that civil violence cannot call for the complete annihilation of identity because political conflict inherently relies on the identification of its victims. In this, the body becomes an artifact that is indispensable to the victorious side, existing as both a human being and evidence of a battle won.

Testimony

When talking to Erichtho, Lucan shifts the evidentiary function of the corpses from recognition into a testimony, where the dead body becomes a source of information that is unavailable to the living. Lucan introduces this encounter by describing the general mood: "all minds were troubled by presentiments of war; it was plain that the fateful hour, the final outcome, was upon them, and that destiny drew near"(VI. 414-416). Even Sextus Pompey, who was "among the uncertain" (VI. 419), is afraid. Even though battle was approaching, Pompey is less concerned with death itself, but the uncertainty that lies in the future. Therefore, he asks Ericto not for comfort, but for certain knowledge of the upcoming conflict, wanting events deprived of their ability to arrive unexpectedly. Ericto says there are many ways of acquiring the truth through Thessalian witchcraft, but because the battlefield is fresh with recently slain soldiers, "the simplest way is to steal a corpse from the Thessalian field;"(VI. 619-620). She does not choose a high-ranking officer, but rather a body that is still in a condition capable of speaking. This detail shows that truth, in this case, depends less on who the person was and more on whether the body is physically capable of the task at hand.

Testimony conventionally relies on a living witness capable of recounting what has been seen or known. However, in Ericto's case, she chooses a corpse, which is then forced to speak to gain knowledge. Throughout this process, the corpse has no agency or say in the process. It is chosen merely for its usefulness, dragged away, reanimated and forced to speak through a demonic ritual. This shows that the civil war has corrupted not only what is considered true but also the means through which the truth is extracted. The witness is given no choice, and evidence is forced out of it, the same way Sulla's men forced identity out of the victim's bodies.

Lucan describes the savage process: "Oh wretched ghost, iniquitously robbed of death's final gift, that is: to die no more! " (VI. 725-726). Even in death, the soul is not allowed to rest peacefully, solely so that information can be gained that should not be accessible in the first place. In Sulla's definition of Rome, the corpse acts as a means of identification, but now, the corpse acts as a channel to knowledge that could not be gained in any other manner. The corpse's testimony is not naturally given, but manufactured through manipulation. Thus, the body no longer serves as a reliable witness, but is used by the living to justify a constructed version of the truth.

Proof

Upon the death of Pompey by one who had previously served him, Lucan notes that "the onset of death had brought no change to the look and countenance of the hero" (VIII.667-668). This attention to Pompey's face does two things. First, it establishes Pompey's identity as Lucan's precise descriptions leaves no room for uncertainty that the man slain was indeed the general Pompey, offering evidence of his death without possibility of mistake. Secondly, it sets up a contradiction because a face precise enough to confirm identity is also precise enough to be worth altering.

Even after death, there followed a long process by which Pompey's head was embalmed and preserved. Lucan describes the process: "the blood was drained from the flesh, the brain removed, the skin dried, the moisture causing corruption was drawn from the innermost parts and, by infusions of drugs, the head embalmed" (VIII. 694-697). This preservation ensures the evidence survives long enough to serve its purpose, since a decomposed head may not be definitive evidence of Pompey's death. The murderers must therefore preserve the appearance of the man while destroying the man himself. This is another example of the recognizable mutilation paradox, except now the mutilation occurs beneath the surface instead of on it, invisible to the person who eventually sees it.

To Caesar, a messenger is not sufficient to prove Pompey's death—yet when he is finally presented with the head, he does not react with the satisfaction one might expect. Instead, he orders it removed from his sight, furious that he has been "robbed of the sole privilege of civil war, that of granting life to a man defeated" (IX.1066-1068). Caesar had wanted to defeat Pompey before pardoning him, a display of clemency that would allow him to assert his authority better than through a severed head received by strangers. The head confirms that Pompey is dead, but it cannot serve the political purpose Caesar needed. Lucan also notes that no

one believed Caesar's grief, "nor did the onlookers believe his grief" (IX.1107-1108), suggesting that his mourning is itself a performance, meaning both the evidence and the reaction to it are manufactured. This reflects the pattern in Lucan's writing that what is seen is not satisfactory, because the same manipulation used to produce evidence can also be used to fabricate a response to it. The civil war is not merely a contest over territory (Caesar had long taken Rome), but a contest over the deaths of specific rivals, and here even the winner's reaction to that death is staged. Caesar's fury is not proof that visibility succeeds where language fails, but evidence that visible "proof" remains just as vulnerable to political manufacture as everything else in the poem.

Yet the head that appears to provide irrefutable evidence is itself a construction. It displays Pompey's features accurately, but everything beneath the surface has been severed, drained, and chemically preserved. Producing this "perfect" evidence required more manipulation than even acquiring the corpse in the first place, meaning the object gains credibility only through the same tampering that should disqualify it as legitimate evidence. This is structurally the same problem as Ericto's corpse as both cases require a body to produce a specific truth outside of its own agency and both cases require manufacturing of a truth that undermines its authority. If evidence must be constructed to be legible, manipulation for the sake of legitimacy may actually decrease it.

As seen through the preservation of Pompey's head , it is obvious that the more desperately civil war demands new ways of certainty, the more violent bodies are manipulated to produce it. Each attempt to manufacture better evidence—Sulla's mutilation, Ericto's resurrection, and the Egyptian's embalming—destroys part of the trust that legitimizes evidence. Pompey's head appears to offer a certainty that political language cannot provide, but its manufacture reveals that material truth can be politically fabricated just like rhetoric can.

Inscription

After physical identity has been fragmented far beyond recovery, Lucan turns from the body itself to marks, names, landscapes, and ultimately his own poetry as methods of preserving truth. When Cordus attempts to bury Pompey's body, it appears lackluster given the glorious funeral an esteemed general like Pompey deserved. Lucan describes this: "Here instead stands a little stele, with no titles, no rolls of office, and Pompey's name that men might read high on the temples of the gods, and over arches decorated with enemy spoils, is barely raised above the sands, so low strangers must stoop to read, and travellers from Rome would pass by were it not made known"(VIII.816-822). He therefore marks the location and writes "Pompey lies here." This shows a progression in how Pompey's body has been used as evidence throughout the poem: first, his face identified him in the midst of the conflict; then his embalmed head served as proof of his defeat; finally, writing identifies his cremated remains. Each part requires more external interpretation than the last, making the body's raw physical form become less useful. This progression shows a shift away from the body's own authority, where Sulla's victims and Pompey's head still relied on physical recognition, however manufactured. A name carved in stone still requires no body at all to function as proof.

Lucan immediately rejects the idea that Pompey and his life's achievements can be contained by a singular stone marking his grave. Pompey's significance extends across the Roman world rather than remaining fixed on the Egyptian coast, and a grave marker, unlike a corpse, cannot claim the physical certainty that made earlier evidence compelling. The claim it makes is asserted rather than demonstrated, and is therefore unconnected to the body that it names. What then emerges is a competition between Pompey's corpse, grave, name, legacy, and finally Lucan's own poetry as means of preserving the "real" Pompey. This is where Lucan comes in; through his poem *Pharsalia*, he fills in the gaps left by an inadequate corpse and grave by reconstructing Pompey's historical meaning himself, replacing failed physical evidence.

In the aftermath of the war, corpses are shown to be scattered all across Thessaly. Lucan describes their remains as becoming part of the earth itself, ensuring that future activities like farming will strike the bones of Romans. Where the poem began by locating evidence in a single body, the war's end expands this evidence across an entire landscape. Civil war has literally written itself into the earth, producing a material historical record that is far harder for any single actor or group to suppress. This evidence has essentially escaped human control, unlike the corpse, which could still be hidden, embalmed, or misidentified.

Conclusion

Lucan's dead cannot simply remain dead, because the living world has lost the institutions that once produced truth on their own. The dead must take on this duty and perform the work once left to political institutions such as establishing identity, confirming facts, and preserving memory. Yet the body is a fragile tool for this as civil war mutilates, decomposes, and reburies it faster than it can be used as evidence. This is why Lucan cannot stop at the corpse. Bodies decompose, faces decay, graves collect dirt and get forgotten, landscapes shift over time. Every material form of evidence in this poem eventually fails in the same way the political language it replaced did. However, poetry does not have this weakness. Unlike a corpse or a grave marker, a poem can be copied and read again without depending on a single physical object being intact.

The body may be the most immediate witness to civil war, offering evidence no institution could provide, but it is also the least durable. Lucan's poetry inherits the evidentiary burden the corpse could no longer carry and in doing so, it becomes the best form of evidence present in the poem. A body made of language, that is no longer subject to the mutilation that defines every other kind of proof in *Pharsalia*.

Works Cited

Dinter, Martin T. Anatomizing Civil War: Studies in Lucan's Epic Technique. University of Michigan Press, 2013. https://doi.org/10.3998/mpub.4586464.

Lucan. The Civil War (Pharsalia). Translated by A.S. Kline, 2014.

McClellan, Andrew M. Abused Bodies in Roman Epic. Cambridge University Press, 2019. https://doi.org/10.1017/9781108697118.

Mebane, Julia. "Pompey's Head and the Body Politic in Lucan's De bello civili." Transactions of the American Philological Association, vol. 146, no. 1, 2016, pp. 191–215. https://doi.org/10.1353/apa.2016.0003

Factors of Happiness in Various Cultures By Melia Outman

Abstract

On the surface, happiness appears to be a common desire, but its root and how it varies across cultures are difficult to define. Studies show that culture is what shapes one's happiness, and within cultures, there are various factors that shape personal happiness. Research shows that happiness can be influenced by factors from personality to income, cultural background to social status, which explains why happiness is obtained and pursued in different ways across cultures. An individual has the freedom to choose whether their happiness is determined by their well-being, success, or surrounding social support. By prioritizing a specific set of factors, cultures confine themselves to a particular definition of happiness that may not always be beneficial in the long term.

Through a deep dive into the research surrounding happiness in Finland and the United States, this paper shows a marked contrast between the two countries. Finland has been declared the world's happiest country for many years consecutively, while the US was rarely near the top. This paper explores where the differences in perception lie by examining the key factors that are used to define happiness in each place. It is not a black and white issue, meaning one country is happy while another is unhappy. Rather, the determining factors are shaped by culture. Finland clearly put an emphasis on their community and political structure and the prioritization of people's well-being, which resulted in citizens trusting their government to provide for basic needs. On the contrary, the US culture pursues happiness based on an individual's success, but while getting lost in the competition, resulting in comparison and jealousy. The Finns reflected happiness that felt natural to them, while the US consistently strives to obtain its happiness through work or achievement.

Overall, this study of happiness shows how many people use it to measure their own worth and value, but looking back, this actually gets in the way of real happiness, because focusing on reaching the top of social status can take away from the true joys of life. This research really hit home for me because I've experienced this contrast firsthand, being raised in a two-culture household, where priorities for happiness looked very different. I've also experienced this on my mission trip, where I noticed the different ways cultures across the world seek fulfillment and joy in their lives. These cross-cultural experiences got me curious and led to my interest in anthropology, which eventually sparked the idea for this paper. In a world that's still facing social, economic, and mental health challenges, it's more important now than ever to explore these differences between our countries and figure out how to create a happy, healthy balance in life.

Differing Lifestyles Across Cultures

Although happiness is considered a universal desire across the world, culture is a determining factor for what values we pursue, which have been gained over the course of life, in order to obtain happiness. While happiness is widely considered a lifetime goal, in reality,

happiness holds different meanings to each person based on the cultural context. Depending on the values being obtained, individuals take different cultural pathways. Across national contexts, these differing levels of happiness are determined by culturally specific lifestyles and value systems. Individual experiences of happiness are guided by the surrounding cultural values, beliefs, and social structures that form the culture and values in which a country is based. It is essential to examine one's own country from an outside perspective to recognize which values are being prioritized compared to the values that are overlooked. The pursuit of wealth, high social status, marriage, freedom, and life satisfaction are treated as critical values in the pursuit of happiness. Research indicates that by stressing one of these, the other priorities receive less attention unless there is a strong emphasis to have a balance of everything. By having a healthy balance of the determining factors of happiness — meaning not putting one value over the other — a person's lifestyle is affected in a big way, creating a healthier lifestyle that adds to their overall well-being. That's why having a healthy balance matters so much, and it's also why a positive culture and lifestyle ends up being reflected in a person's happiness across different cultures.

In everyday contexts, the people and environment we surround ourselves with will affect how we act. Eventually, we begin to adopt the standards based on the surrounding environment, which cultivates our experiences, directly translating into why humans don't operate in a universalistic world. This supports the argument that “any adequate understanding of seemingly universal factors can never be complete without taking into account culture-dependent ways in which such factors are realized and allowed to shape happiness and well-being” (Uchida et al. 223). In other words, although individuals may value many of the same aspects of life, culture plays a significant role in determining how those values are interpreted, prioritized, and experienced. Since the life-altering Covid-19 pandemic in 2020, the pattern shift has shown an increased preference for a more comfortable, content way of living. Contentment can be viewed in many different ways. One way is being happy with the bare minimum, and another is working intensely to a point where you are satisfied with your life because you have obtained more than necessary.

These contrasting lifestyles are directed by mentality and can also vary based on personality. On the other hand, both of these ways can bring happiness, depending on the person, explaining why some individuals are satisfied with less while others are not. It tends to be harder to find contentment amongst materialistic societies due to people viewing “Possessions as a sign of success and consumers tend to display their possessions because they believe that material goods help them make a positive impression on others as creating happiness” (Eternod 26). By doing so, people who prioritize their materialistic possessions primarily focus on how they are perceived while using the idea of "creating happiness” for others as a cover-up for their hidden desire for social recognition. The concept behind the idea of physical things “never being enough” comes from materialistic items due to comparison. Social comparison encourages the focus to be on what hasn't been achieved instead of being appreciative of all that has already been obtained. “Possessions are seen as a sign of success and consumers tend to display their

possessions because they believe that material goods help them make a positive impression on others as creating happiness, and as being central to their lives" (Eternos 26).

People have claimed that they prioritize materialistic possessions solely to spread happiness, but this mindset ends up shifting sight of current existing sources of fulfillment and instead letting comparison take control over one's life. Just as individuals may use a mask to conceal and justify their actions, this idea is similarly applied to the relation between happiness and beauty. In the United States, cultural expectations often place significant value on both happiness and physical appearance, leading some individuals to project an image of success and fulfillment in order to fit in with whatever standards society expects. For example, makeup can create the illusion of beauty, just as wealth or social status are displayed to give the perception of being happy. Expressive suppression is defined as "the inhibition of emotional display and has been consistently linked with poorer psychological health" (Xu, Zhang, Tsai 1). This way of living can often lead to depression or a bad mental health state, suggesting the importance of social environments in which individuals should be allowed to express their emotions.

This paper will dissect differing views on the factors contributing to the obtained happiness and then compare two drastically different countries that reflect contrasting happiness standards. The two countries that will be used are Finland and the US to provide specific information on how the overall happiness status is determined by their own country's desires, such as freedom, satisfaction, or national pride. According to the 2019 Happiness Report, "Finland is the happiest country in the world, with Denmark, Norway, Iceland, and the Netherlands holding the next top positions" (Tenhunen 18). In contrast, the US also reports high levels of happiness, but they are grounded in different cultural values. This demonstrates how the prioritization of various factors influences well-being, as Finland and the United States emphasize different cultural values that shape distinct social norms and understandings of a fulfilling life within their societies.

One Experience Opening a Perspective

Before I get into my own story, I want to be upfront about what this part of the paper actually is. Most of what I'm using here comes from other people's research on happiness across cultures — psychology and sociology studies, plus data from the World Happiness Report — combined with my own personal experience from growing up in two family cultures and going on a short mission trip to Uganda. I didn't do interviews, surveys, or any kind of formal research of my own. So the personal stories below should be read as just that: my own reflections that got me interested in this topic in the first place, not proof of what life is actually like for everyone in Uganda, Finland, or America.

Personal experiences of mine will illustrate what it looks like to take negative experiences and then shift to a positive environment despite circumstances. During a mission trip to Uganda, it became evident that most of those people had to make the most out of what they had because they already lived in not the best physical state. The level of visible, authentic happiness in an environment surrounded by constant setbacks was particularly striking. Despite

their poor conditions, their strong presence of joy and social connection was still present, which prompted the curiosity to further examine how such happiness was possible. After being exposed to this new culture, I was encouraged to adopt this positive outlook on life. Stepping from an outside perspective allowed me to see the differences. While kids at my school complain about traffic, the Ugandan children do not complain about 2-3 hours of walking each way. I have noticed a striking contrast: even though the US offers naturally better living conditions, the people I met in Uganda possessed a genuinely more positive and resilient outlook on life. It eventually came to light that the Ugandans preferred to prioritize tradition and a tight-knit community, thus explaining why everyone there kept positive attitudes most of the time. This experience drew parallels to my own family back at home.

Growing up in both Mexican and White family environments exposed me to two distinct cultural value systems. The contrast is apparent when we have family get-togethers with my mom's side of the family compared to my dad's side. Gatherings on the paternal white side tend to feel very organized, planned, and like a contained environment, which reflects the values of how each member was raised since childhood to eventually form this environment. Conversely, family gatherings on the maternal Mexican side were more like celebrations where family connection was clearly emphasized. The contrast was a strong reflection of how countries can adopt similar values influenced by whichever factors a country chooses to pursue. Overall, the factors that shape happiness vary across countries, which explains why happiness differs across cultures. By comparing Finland and the United States, it becomes clear that what a society chooses to prioritize does not just influence individual happiness, but ultimately defines the values, expectations, and overall character of the society itself.

The Origin of True Happiness

Anthropologists have recognized that, while happiness is a universal desire, individuals interpret its foundation differently despite the natural drive to pursue happiness (Eternod; Veenhoven; Diener & Lucas). This understanding eventually led to the realization that happiness is influenced by both the country in which one resides and the surrounding environment. So anthropologists have observed that the expression of happiness varies across countries. How happiness shows up depends on the cultural context, explaining why there are contrasting views on how happiness is achieved. According to most experts, happiness is defined "as subjective well-being and satisfaction [which] is commonly used to conceptualize happiness" (Eternod 7). Cultural anthropologists use the idea of subjective well-being, the term for when someone has their personal self-assessment of their own happiness, to explain why they think that is how happiness should be defined. It is the same idea that is derived from cognitive thinking, which is the human physiological drive to seek positive things in an environment. Described in contemporary literature, "subjective well-being is typically defined as an overall cognitive appraisal of the quality of one's own life". This refers to one's evaluation of their life as a whole, supporting the concept that happiness should be measured through satisfaction. True happiness is an emotional state that develops as a result of life satisfaction, rather than something that can

be physically ordained. Satisfaction can come in many different forms, whether you become satisfied through achieving wealth, a healthy marriage, or even things such as popularity. The opposite of cognitive or cognition would be emotional, which is described as the intense emotions that occur when expressing contentment. A paper by Eternod mentions “the key to happiness can exploit our full potential, not too focused on the pursuit of happiness, [but] control our own lives, use material things to our well-being, and seek satisfaction in our life” (Eternod 7). It is implied not to strive for happiness and make it the sole purpose, but rather to take control of life and seek satisfaction instead. Satisfaction is actually derived from the feeling of pleasure or fulfillment of one's wishes, expectations, or needs, and this all leads to our state of happiness.

Unlike the first perspective, the second view draws the line between happiness and satisfaction. While one expresses how important it is not to chase happiness, it is actually supported by the other view, explaining how happiness will come along with seeking that satisfaction first. At times, the pursuit of short-term happiness and satisfaction can risk long-term happiness. Veenhoven supports this idea by stating that “happiness roots in the gratification of basic needs that are part of human nature” (5). Though our satisfaction can come from our needs being fulfilled, it is commonly confused with people’s wants because both have the ability to be fulfilled and bring happiness. Needs are described as biological because they are a necessity to life, which is why they are a part of human nature. Wants are a part of human experience, which explains the reasoning behind why wants are cultural.

While the capacity to want is biologically a part of human need, what specifically is wanted is shaped by our environment, society, and culture that comes with life experience. The tendency to want what others have, commonly driven by comparison and envy, is heavily influenced by cultural surroundings. Cultural surroundings determine what is valuable, prestigious, or necessary, ultimately shaping the entire structure of a society. In order to develop wants, you have to live life a bit to discover what you want, but before that develops, needs are a necessity first. On the contrary, wants are typically something you would find on a wish-list because they are nice to have but are not necessary to living. Having wants is a privilege, and that is why you will not find people making their wants a priority because they have yet to achieve their needs. A society that is primarily focused on wants being fulfilled, like a sole materialistic environment, can result in an unhappy society.

What are These Factors, and Who are They Affecting?

Maslow's Hierarchy of Needs Theory depicts basic human needs that are organized into a hierarchy. The hierarchy consists of various levels, ranging from physiological needs to internal growth needs. At the bottom of the pyramid lie physiological and safety needs, while at the upper part of the pyramid are self-actualization, esteem, love, and belongingness. The theory claims that the needs at the bottom of the pyramid should be prioritized and cared for first because it is necessary to satisfy these needs before approaching fulfillment in internal, self-addressing needs. Different countries represent fulfilling this theory depending on where their happiness has been built upon and where the foundation lies. From Veenhoven's perspective, he also claims, “If

happiness depends on the degree to which human needs are met, average happiness will be higher in societies that fit human nature better than others" (Veenhoven 12).

Ultimately, both cultural norms and geographic location play a big role in shaping social outcomes. If that environment allows your needs to be met, then you would find yourself in an overall happy community, which results in better living and creates room to actually have those wants. Surrounding yourself with a healthy community will open up opportunities that are separate from needs, specifically the things considered a privilege, like the "extra" desires in life. Ultimately, the extent to which a society enables the fulfillment of human needs shapes collective well-being, determining if individuals can move towards culturally defined wants. This relationship reinforces how happiness is not only found within an individual's emotional state, but is also shaped by societal structures like government policies, cultural values, economic stability, and societal expectations.

There is an emphasis on the different types of groups that are affected by the different values and factors of happiness. As a human experiences life, enough different positive and negative experiences have taken place to make all the difference in how an individual views life or perceives happiness. Positive experiences consist of the perception of living standards, personal freedoms, or social networks. It is shown in the 2019 Global Emotions Report that "Latin American countries lead in the positive experiences ranking" (Eternod 23). This could mean that the Latin American countries are consistently seen as laughing, smiling, feeling respected, and showing enjoyment. Mexico's positive experiences surround social support, which is a contributing factor to their happiness. Among Latin American countries, people with a high perception of social support are more satisfied with their lives because they rely on their tight-knit community when they are in their hardest moments. It is said that "cultural values might determine both the extent to which positive experiences are likely to be sought out and experienced and the extent to which they are likely to be impacting overall happiness" (Eternod 33). The statement breaks down two core influences, suggesting that an individual's cultural background acts as a filter for understanding the world around them.

By focusing on how people seek and experience life, it illustrates how deeply cultural norms dictate the pursuits we value. Secondly, cultures determine how a specific moment carries into long-term well-being thus impacting overall happiness. Negative experiences are the opposite, they consist of health problems, sadness, and stress. "On the other side, at the top of the negative experiences are ranked Africa and the Middle East which are experiencing severe conflicts" (Eternod 24). This data is contradictory because it misinterprets regional well-being; the lower scores in Africa and the Middle East do not reflect a lack of happiness, but rather show how cultural context shapes life experiences and opportunities. Rather, it is even more challenging to stay happy with a positive mindset while living in negative experiences of life, experiencing things such as poverty, abuse, abandonment, or loneliness. Veenhoven's work shows that "The overall appraisal of life draws less on cognitive comparison with cultural standards of the good life" (Veenhoven 13). When people evaluate how comfortable they are with their own lives, it results in relying more on daily experiences and personal feelings rather

than how an individual's life matches the ideal standards defined by cultures. Overall, cultural values dictate how people interpret their living conditions and how group experiences affect their personal happiness.

Happiness is not a random occurrence but a result of specific foundational factors. This pursuit of happiness calls for "understanding of conditions for happiness" (Veenhoven 1). To create the best conditions, you must build an environment where people's needs are met. Primarily, Veenhoven's research found that the main factors affecting happiness are freedom, social rank, personality, married status, income, satisfaction, national identity, and comparison.

With a Factor Comes a Purpose

Freedom is defined as "the power or right to act, speak, or think as one wants without hindrance or restraint." (Oxford Languages) Therefore, when individuals have freedom, they feel like they have control over their lives, resulting in increased happiness. This translates similarly to how kids desire freedom as they get older because it is a big part of growing up. If social ranking was demonstrated on a "ladder," then successful individuals would be found at the top, and people below are either comfortable at that level or are trying to work their way up to the top as well. "People are typically happier on the upper steps of the social ladder than at the bottom. This appears in findings on relative income position, occupational prestige, subjective class identification and indexes of socio-economic status" (Veenhoven 8). Though it depends on the situation, personal experience contradicts this statement due to happiness still being demonstrated even when surrounding conditions do not provide the ability to get anywhere near the top of the ladder. Therefore, even when found at the bottom of the ladder, happiness is still possible, but studies show that it is more common when residing at the top.

According to personality, extroverts tend to be happier compared to introverts. Studies have shown that there is a universal link between extraversion and hedonic level. Extraversion is a trait characterized by high levels of energy, talkativeness, and sociability. Hedonic level relates directly to extraversion because it is when an individual experiences reflect positive emotions and fewer negative emotions. Individualistic societies tend to have a stronger sense of self-esteem than collectivistic societies, though that's not true across the board — there are plenty of exceptions. While married adults generally report higher levels of happiness, this correlation varies significantly based on income, gender, age, and individual personality traits. This correlates with the contrast between extroverts and introverts, as introverted individuals may experience equal levels of happiness when living alone.

It is said that "money can buy life satisfaction and soften life's adversities but can't buy happiness" (Eternod 26). Income impacts cognitive and emotional components. Cognitive well-being is positively linked to wealth, as emotional happiness tends to rise with higher income. Income also affects where one stands on the "social rank ladder" because a higher income improves material living conditions, but it also increases pretensions and expectations. Therefore, such conditions transform because they begin to get neutralized by a change in needs.

This can affect happiness in different ways because there are also areas with little to no material living conditions, yet they remain fulfilled and satisfied.

A high satisfaction level in one area can spill over into another domain and then boost overall subjective well-being. Both life satisfaction and global evaluation are significantly influenced by national identity. National identity is defined as “a person's sense of belonging to a specific nation or state,” (Oxford Languages) which has been shaped by the surrounding culture, history, language, and values. Most people have a desire to belong, so when those wants are met, it increases one's overall happiness. On the contrary, comparison limits happiness to individuals who have a strong tendency to compare their situation to others. This negative cognitive habit inhibits well-being, which explains why it is less commonly observed among extroverts. Some are born without the option and are forced to live underprivileged. Those conditions result in the feeling of helplessness and powerlessness at the mercy of uncontrollable external forces.

Overall, the course of an individual's life is determined by the factors one chooses to live by. The cultural values stated can shape individuals' own concept of happiness and redefine how it is experienced. Veenhoven describes happiness as the “overall appreciation of one’s life as-a-whole” (Veenhoven 2). Happiness does not necessarily require complete satisfaction with every aspect of life; rather, it can be experienced through fulfillment in areas that individuals personally value most. This paper defines happiness as a natural, positive feeling that can be obtained when reaching a personalized satisfaction based on cultural normative standards, such as the determining factors stated above.

The Root and Map to Finland's Aspiring Happiness

According to the World Happiness Report, Finland has been declared the happiest country in the world for the last eight years, followed closely by Denmark and Iceland. The unique combination of factors driving Finland's high quality of life is rarely replicated in other nations, explaining why the country stands apart in global happiness rankings. The five Nordic states: Finland, Denmark, Norway, Sweden, and Iceland, have all been in the top 10 happiness rankings, so what is special to this region of countries? “Nordic citizens experience a high sense of autonomy and freedom playing an important role in determining life satisfaction as well as high levels of social trust towards each other” (Tenhunen 22).

The Nordic high life satisfaction is based on two things that reinforce each other: individual empowerment and community stability. Having a high social support in a county will bring people together to form a culture and a tight-knit community. Their trust in institutions builds foundations and trust due to a welfare state model with social benefits such as free education, stability, protection, housing allowances, or child benefits. If state institutions deliver what they promise to their people, it gives off the perception of generosity in taking care of their citizens, amongst adversity, which is exactly what Finland exceeds. The Finns feel cared for and acknowledged, adding to their happiness levels. Finland's happiness is due to factors surrounding institution quality, reliable welfare benefits, low corruption, freedom in decisions, social support, and a well-functioning government.

Finland emphasizes equality, which separates it from other countries. Finland takes pride in its government system; their citizens trust their government, resulting in more agreement and a strong sense of community. It functions exceptionally well due to an organized government that mirrors the way society is structured, which makes people trust their institutions even more. For example, Finns do not have to worry about homelessness because their institutions make sure that there are available housing options. This strategy is viewed as less homelessness equals less corruption, creating a safer environment compared to most. Also, Finland believes that having a fair schooling system at equitable schools for kids will set them up for success in the future, providing Finland with a strong foundation to build off of for future generations. Money problems are usually one of the biggest sources of stress for people, so getting rid of healthcare costs in Finland doesn't just keep people physically healthy — it also makes a huge difference for their mental well-being.

Unlike many countries, Finland maintains a largely government-funded healthcare system that reduces financial burdens and simplifies access to care, creating a more secure and less stressful environment for its citizens. It is hard to obtain spots in a hospital and, in actuality, takes up to hours to meet with a patient. Rather than prioritizing the pursuit of wealth, Finland emphasizes economic stability, social support, and equality. Strong public institutions and a solid welfare system help ensure that citizens' basic needs are met, reducing financial insecurity and contributing to higher levels of trust, well-being, and life satisfaction. Because economic inequality is comparatively lower, many citizens experience greater economic security and social trust. Finland's success in promoting well-being does not mean that every citizen is happy, nor does it suggest that happiness is unique to the Nordic countries. Although Finland's emphasis on well-being existed long before the COVID-19 pandemic, the crisis reinforced the importance of social support, stability, and sustainable public institutions by highlighting citizens' vulnerability during times of uncertainty.

Since then, Finland has begun prioritizing strengthening their work-life balance, mental health, and long-term happiness. To summarize, Finland has a sense of community and unity prompted by their “good governance, equality, excellent educational attainment and very little corruption on a global scale,” which are rare qualities that reduce overall stress and instead foster greater overall societal contentment, explaining why it is consistently ranked the happiest country (Tenhunen 25). Finland highlights how national identity satisfaction, freedom, and social support drive happiness specific to the Nordic countries, while factors such as income, marriage, and personality play smaller roles. This demonstrates that happiness is shaped not only by individual desires but also by the cultural and societal structures that surround them.

The United States Takes a Different Road to Happiness

In contrast to Finland, happiness in the United States has a lot to do with income, especially considering it is a highly consumeristic society. A consumeristic society is defined as “In which the entire society is organized around the consumption and display of commodities through which individuals gain prestige, identity, and standing” (Firat et al.). “Much of the

observed positive association of income and happiness in the United States can be accounted for by the association of income with health, education, age, marital status, unemployment, and race" (Social Trends In American Life). In the US, income is strongly prioritized by citizens to the point where it is less seen as a factor, but more like a common desire. Some even prioritize income over happiness or believe that striving for income primarily will bring happiness. Although many people associate wealth with happiness, this connection is not universal. In a culture that places a strong emphasis on financial success, individuals may come to view income as a key indicator of well-being.

However, people with wealth can still be sad. Another philosopher said, "For Americans at a given level of income, those whose age peers enjoy higher incomes tend to be less happy than those whose age peers are poorer" (Social Trends In American Life 280). As wealth increases, physical possessions are gained, and status is raised, making it easy to compare situations to those of others. It is an added pressure to US citizens' lives to meet the surrounding standards and possibly match incomes. This is what makes the US competitive, but as the focus has shifted to wealth, the desire has been less on happiness, which explains why it is decreasing. The US is not entirely negative when considering separate factors besides wealth, such as job satisfaction and the purpose that comes with that. Unlike Finland, income became the primary focus for the United States, and while it is not entirely good, money can make some things easier, especially in the US. The concept of income also affected external factors because once it was seen how money is valued and strived for, suddenly, education or healthcare were not affordable to the average person anymore. Schooling and health are simple necessities of life, but the economy saw it as an opportunity to grow and expand by charging for the better levels of service. It encourages people to want more money, and naturally, most humans want the best of what is being offered to them to reflect success and the ability to afford certain things.

Though social hierarchies have taken different forms, the idea of social status draws a border amongst the citizens, creating greater division and less connection. This translates as to why the United States was perceived as a country that surrounded individual success, implying that the US is not very community-based, which is the opposite of Finland. Lawrence, Rogers, and Wadsworth declared that "Higher levels of subjective well-being are also related to increased success in a variety of domains, including social relations and economic stability" (Lawrence 2). Lawrence supports this by relating subjective well-being to success, but it also raises levels of stress because it brings added pressure to meet requirements. Stress inhibits happiness because it is difficult to be stressed and happy at the same time. While Finland emphasizes social support and increases their happiness, the US's individualistic culture has led to isolation, resulting in declining happiness. The contrast demonstrates the importance of a country's culture because it reflects on well-being and character.

Can Finland and the US Represent Happiness on the Same Level?

"The American dream is known to the whole world, but Finns may not know the Finnish dream themselves" (Tenhunen 26). The American dream is a conceptual idea perceived to be

focused on individual ambition and going from “rags to riches.” On the contrary, the “Finnish dream” is viewed as a quiet, structural lifestyle that has been deeply incorporated into Finland to the point where it is what they are known for and reflects opposing values compared to the US. The problem lies within what each country prioritizes because both the United States and Finland think differently about what will bring happiness. While the US strives to be recognized for its income and success, the Finns prioritize mental health and well-being before anything else. Some have claimed it is better to be healthy than rich, but the way one's mindset tends to be influenced depends on the surrounding environment and what an individual is looking for in life.

Lawrence, Rogers, and Wadsworth argue that “Happy people typically enjoy better health outcomes because they demonstrate more successful adaptation; better problem-solving skills and coping strategies; more creative, imaginative, and integrative thinking; greater resilience; and a greater ability to deal with adversity” (Lawrence 2). This quote provides a strong foundation for framing happiness and, therefore, depicts all the positive effects, showing that happiness isn't just some fleeting emotion — it actually helps people adapt and cope with life. Happiness results in one's health in a positive light, bringing out joy, and is reflected through people's actions. Happiness specifically improves mental and physical health, life outcomes such as productivity, relationships, and creativity.

Existing in both Finland and the US is national pride, which is the emotional attachment to one's own country, but it is conceptually different from national identity. “On the one hand, national pride can be based on positive feelings toward a nation’s accomplishments. On the other hand, national pride can represent beliefs about superiority of one’s own country, based on positive feelings toward its characteristics” (Ha, Jang, 473). National pride is rooted in Finland as internalized, but in the United States as expressive. Finland's structure and organization speak for themselves, as they do not need to express their pride because it is already universally seen. Finland’s pride is focused on the functioning success to contribute to their state as a whole to improve what it once was. Finnish citizens are proud of how far their country has progressed, particularly through the development of its social welfare system and public services.

On the contrary, American pride is solely individualistic but expressed very loudly to demonstrate their strong values in freedom, which separate Americans from many other countries. Finland identifies with their shared values of quality and education, contributing to what makes it a happy country. The US identity is built based on liberty and personal achievement because they surround less collective values. Personal success directly leads to individual happiness, but not the country as a whole, like Finland. “Individual countries have already made happiness an attribute of state well-being” for themselves; rather, it will not lead to a whole country's well-being (Tenhunen 27). Happiness within a society is not individually based whatsoever because the World Happiness Rankings are determined by country and not person.

Markku Ojanen, Finnish professor, describes individual happiness “as a very holistic space, with other emotions: joy, satisfaction, love, gratitude and hope. Equally, people seek meaning, inner peace, security, love and self-worth in their life” (Tenhunen 19). Certain countries are more populated compared to others due to where an individual's qualities are valued.

Talented, innovative people are attracted to areas that will recognize them for their qualities. Societies that don't attract this kind of talent risk falling behind, both in security and progress. Finally, there is a distinct difference between when happiness begins to create a society versus when happiness transforms a society. It highlights the critical distinction between the foundational role of individual contentment and the structural shift caused by a collective change.

Key Applicable Take-Aways

Based on what has been demonstrated by Finland, for a country to achieve happiness, it is necessary to elect a government that is respected and will provide equality, along with a certain degree of freedom, efficient schooling, housing, and healthcare. These are all the factors that contributed to Finland's happiness, making it the number one happiest country in the last 8 years. The trust factor is huge because it doesn't come easily, but takes time for a country to build. Nevertheless, happiness is not permanent or guaranteed to last indefinitely, but it is always what you make it and will alter personal well-being, and it ends up affecting other parts of life too. When individuals actively cultivate their own happiness, others observe this behavior and naturally emulate the positive traits being demonstrated. "Higher levels of subjective wellbeing are also related to increased success in a variety of domains, including social relations and economic stability" (Lawrence 2). This reflects the core finding that happiness is nowhere near a consequence of success, but it's more of a cause. Success often brings happiness, but sometimes successful people still end up lonely because their whole life was focused on getting to the top and nothing else.

Social support has consistently been seen among happier people because when people go through their darkest moments, falling into depression is likely, which is when it is necessary to have surrounding support to ensure this does not take place. The traditional economic terms for happiness surround material wealth and an individual's rational choices. Always aiming for a higher income is one of the objectives that has been commonly seen in the United States, but it is not the main measurement used to determine quality of life. It implies that societies should always be striving for happiness instead of economic growth because once unselfishness is demonstrated, and the prioritization of a country takes root, then happiness that has been inherited shall naturally increase economic growth. "Various key institutional and cultural indicators of good society feed into each other including well-functioning democracy, generous and effective social welfare benefits, low levels of crime and corruption, and satisfied citizens who feel free and trust each other." This is the recipe for forming a happy society that has an organized structure to it, just as observed in the Nordic countries. It shows how good institutions and cultural attitudes feed off each other in a cycle, building up to high levels of societal happiness overall. Instead of working alone, these factors basically keep each other going — trust, efficiency, and fairness all feeding into one another — which is exactly what happens in countries like Finland.

Limitations

This analysis has some real limits I want to own up to. First, I'm working from a pretty small set of sources instead of doing my own original research, so what I'm saying here is more of a summary of what other people have already found, not brand-new evidence. Second, my Uganda story is just one trip and one experience — I don't want it to sound like poor communities are automatically happier, because that's an idea people get wrong a lot when they visit somewhere different from home for a short time. Third, a lot of the happiness data I used, like the World Happiness Report, is based on people rating their own happiness, which can be affected by culture and by what people feel comfortable admitting. And finally, since I didn't do any interviews or observe people firsthand, everything I say about what life is actually like in Finland, the US, or Uganda should be taken as my own interpretation, not as proven fact.

A World Full of Universal Desire

To conclude, happiness is a universal desire, but the way people achieve it often depends on the culture and values being pursued. Different environments are influenced by the factors of wealth, marriage, comparison, satisfaction, freedom, personality, and social rank, all creating various countries and individual lifestyles. What individuals pursue reflects their ultimate goal. Depending on whether citizens want to live a fulfilling life surrounded by a community of support or be satisfied based on income, it will say a lot about the character and culture. Though individuals tend to prioritize a singular factor because it may be easiest to fulfill, having an equal balance in worldly desires is essential to shaping moral values. Balancing out the factors in life will ultimately bring more happiness than one factor could provide. So every region is different, but lately it's gotten easier to just go along with whatever's popular in order to fit the image of a “perfect life,” which has caused material comfort to increase over time.

Beyond these institutional and economic factors, happiness is also sustained by informal social bonds. In recent years, we see less effort put into preserving traditional values and more into self-success, but this idea strays away from what keeps a community tightly-knit together. All humans go through rough patches in life, but it is hardest to get through them alone, which is why having people's support matters so much. The Mexican heritage, from experience, and its love have remained consistent throughout generations, demonstrating the importance of keeping connections because life is meant to be lived with others and not individually.

Even without achieving any of these factors, gratification is always possible for everyone, and it doesn't cost anything. Being grateful for what you need instead of focusing on what you want should just be part of human nature. Through how we express ourselves, we influence others, which is why the simplest mindset shift from “we have to” to “we get to” can make all the difference in how happiness is seen and reflected by others. Moving forward as a nation, taking into consideration how others express their happiness can help us understand the thought behind their actions, which helps avoid conflict and brings the world closer together. Uganda lived this out. Ugandans took others into consideration because they knew everyone was experiencing life for the first time together, so at the end of the day, it's a learning opportunity for all!

Works Cited

Diener, Ed, and Richard E. Lucas. “Explaining Differences in Societal Levels of Happiness: Relative Standards, Need Fulfillment, Culture, and Evaluation Theory.” Journal of Happiness Studies, vol. 1, no. 1, 2000, pp. 41–78.

Eternod, Vivian. “The Cultural Perspective: Are Some Societies Happier Than Others?” Consumer Happiness: Multiple Perspectives, Springer, 2021, pp. 23–41.

Firat, Aytekin, et al. “Consumption, Consumer Culture and Consumer Society.” Journal of Community Positive Practices, no. 1, 2013, pp. 182–203.

“Freedom” and “National Identity.” Oxford Languages, Oxford UP, languages.oup.com.

Ha, Shang E., and Seung-Jin Jang. “National Identity, National Pride, and Happiness: The Case of South Korea.” Social Indicators Research, vol. 121, no. 2, 2015, pp. 471–482.

Lawrence, Elizabeth M., Richard G. Rogers, and Tim Wadsworth. “Happiness and Longevity in the United States.” Social Science & Medicine, vol. 145, 2015, pp. 115–119.

Maslow, Abraham H. “A Theory of Human Motivation.” Psychological Review, vol. 50, no. 4, 1943, pp. 370–396. PsychClassics, psychclassics.yorku.ca/Maslow/motivation.htm.

“Social Trends in American Life: Findings from the General Social Survey Since 1972.” Google Books, books.google.com.

Tenhunen, Marja-Liisa. “Why Finland Is the Happiest Country in the World?” Cogito: Multidisciplinary Research Journal, vol. 12, no. 3, 2020, pp. 18–28.

Uchida, Yukiko, Vinai Norasakkunkit, and Shinobu Kitayama. “Cultural Constructions of Happiness: Theory and Empirical Evidence.” Journal of Happiness Studies, vol. 5, no. 3, 2004, pp. 223–239.

Veenhoven, Ruut. “Does Happiness Differ Across Cultures?” Happiness across Cultures: Views of Happiness and Quality of Life in Non-Western Cultures, Springer, 2012, pp. 451–472.

---. “Happiness: Also Known as ‘Life Satisfaction’ and ‘Subjective Well-Being.’” Handbook of Social Indicators and Quality of Life Research, Springer, 2011, pp. 63–77.

Xu, Y., G. Q. Zhang, and W. Tsai. “Longitudinal Associations Between Expressive Suppression and Psychological Health: The Moderating Role of Authenticity and the Ambivalence over Emotion Expression.” Journal of Counseling Psychology, vol. 73, no. 1, 2026, pp. 116–127.

LA County Medical Professional Perspectives on Physician-Assisted Suicide in Relation to Ethical Theory and International Law By Noa Sarfati

Abstract

Physician-assisted suicide (PAS) is a central point of debate in modern medical ethics. This study explores the extent to which medical professionals in Los Angeles County view PAS as ethically acceptable and compares their perspectives with established ethical theories and selected international legal frameworks. To do so, the researcher used a non-experimental and mixed-methods design. The findings indicated that respondents generally considered PAS acceptable under very specific and regulated conditions, especially in situations where the patient is terminally ill and of cognitive capability to make a critical decision. Respondents usually chose Principlism as their favored ethical framework, with 73.33% identifying autonomy as the most important biomedical principle. Support declined significantly in scenarios where the patient was non-terminal, suggesting that patient autonomy does not take priority over the fact that there may still be available treatment options. Although the limitation of a smaller and concentrated sample prevents a broad generalization, the findings support the hypothesis that clearer ethical and legal guidance in the field of end-of-life care must be established.

Introduction

Medical ethics represent a long-debated issue, dating back to the early times of Hippocrates. For much of history, death was viewed as a definite sentence that came without man being able to delay it. However, medical advances over the last few centuries have drastically changed the way humans are able to face disease, injury, and, consequently, the end of their lives. Polio, which likely would have paralyzed a child in the 1940s, has now been almost fully eradicated with vaccines (World Health Organization). Patients with no brain activity can still be kept alive through ventilators, feeding tubes, and organ support (Nguyen). Though this progress has allowed many to stay breathing, it has also raised many complex ethical questions regarding patient autonomy, quality of life, and the idea that caretakers must do everything in their power not to harm. Controversies regarding end-of-life decision-making, therefore, occupy an integral part of the discussion in medical ethics and law.

Euthanasia is the general term for purposefully ending a life. There are two kinds: active and passive. The distinction between each one is a thin line, but it is usually defined by whether death is caused through direct action or the omission of life-sustaining measures. Passive euthanasia means that the doctor lets death happen by stopping medical intervention. For example, they may turn off life support, thereby exposing the patient to their ongoing fatal condition. In comparison, active euthanasia is done directly by the medical professional, often in the form of a lethal injection (Gesang). Physician-assisted suicide (PAS) is a halfway point between the two: the physician provides the means for death, typically through the prescription of lethal medication, but the patient must take the dose and perform the final act (Spurdle et al.).

Existing literature on PAS has used a variety of ethical frameworks to advocate for or against it, including the consequentialist-based approach of utilitarianism, the duty-based approach of Kantian deontology, and the morality-based approach of virtue ethics (Hursthouse and Pettigrove; Kranak; Singer). Principlism, however, remains at the core of clinical ethics. This framework centers on four key principles: autonomy, beneficence, nonmaleficence, and justice (Bulger). In biomedical ethics, these principles often yield conflicting conclusions on the permissibility of PAS, making it difficult for countries to establish definitive and coherent legislation. Although some countries have made significant progress, such as the Netherlands and Belgium, others, like Japan, have refused to enter the debate entirely, despite their medical professionals advocating for legislation to be created (Asai et al.; Rurup et al.).

This paper aims to answer the following question: *To what extent do LA County medical professionals perceive physician-assisted suicide as ethically acceptable, and how do their views align with ethical principles and the policies of selected countries?*

Literature Review

Prior to understanding the perspectives of physicians, ethical principles, and country policies, it is first important to investigate the basis of these beliefs and the historical context they stem from. Rather than advocating for or against legalization, this study aims to analyze why those at the forefront of the debate hold the opinions they do by comparing them to what has been established in medical ethics as far back as the 5th century B.C. (David Geffen School of Medicine).

The Hippocratic Oath

The Hippocratic Oath has historically been the foundation of Western medicine, originating in ancient Greece as a pledge professionals must make. It emphasizes beneficence, confidentiality, and most notably, nonmaleficence, also known as the promise to “do no harm.” A researcher from the University of Athens, Dr. George Daikos, notes that over the centuries, the Oath has been interpreted in different ways and applied to movements such as humane medicine, which have extended physicians' responsibilities beyond literal treatment to encompass compassionate care (Daikos). Its applicability to contemporary medicine, however, has grown to be increasingly questioned. Two researchers from the Indian Journal of Psychiatry explore these debates. They argue that modern medical developments, like organ transplantation, reproductive medicine, and, most relevant to this paper, end-of-life decision-making, consistently challenge the Hippocratic Oath. PAS is especially controversial as nonmaleficence can be interpreted as either avoiding causing death due to the outcome or preventing prolonged suffering (Indla and Radhika). In this context, the Oath is still symbolic but functions more as a cultural artifact than a practical ethical guide, given that, as Daikos observed, it was never designed to accommodate the medical evolutions that have occurred after its creation.

Arguments for PAS

Arguments in favor of physician-assisted suicide often center around compassion and autonomy. One factor often less acknowledged but argued to be as crucial, according to Dr. Kittay from Stony Brook University, is ethics from the perspective of dependence. Dependence by a person with chronic or disabling conditions is not undignified, but sometimes means having to rely on the morals of others. Consequently, those disabilities question how a just society should ethically acknowledge their necessary reliance. Within this model, PAS can be interpreted as consistent with compassion and respect for the wishes of patients, as it recognizes their right to dignity (Kittay). Oncologist Dr. Emanuel similarly examines the moral justification for PAS. In addition to his emphasis on the importance of the "patients' autonomy to control their own lives," he adds that "beneficence in relieving excruciating pain and suffering" is also a significant part of the argument. The integration of clinical evidence with philosophical reasoning directly challenges the assumption that legalization inevitably leads to misuse. There is a wide misconception regarding the frequency of employment of PAS. Contrary to popular belief, over 99% of Americans do not take these measures. Even in countries where it is fully legalized, like the Netherlands, "more than 96% of all descendants do not have these interventions" (Emanuel). This statistical reality reinforces Kittay's point that legalization, when accompanied by transparent regulation and ethical oversight, does not normalize voluntary dying, thereby challenging the common "slippery slope" objection that allowing PAS would lead to the normalization of voluntary dying.

Arguments against PAS

In contrast, arguments opposing physician-assisted suicide often rely on traditional medical ethics from centuries past, as well as concerns about systematic vulnerability. Historically, the Hippocratic Oath has served as the primary ethical reference point for these objections. This is consistent with the study conducted by researchers Indla and Radhika, who explain that the concept of "do no harm" is interpreted as anything that causes death, regardless of context. Therefore, those with more orthodox ethical beliefs will take it to mean that PAS is never acceptable (Indla and Radhika). Beyond the Hippocratic tradition, structural concerns have emerged around the financial incentives that legalization may create. As noted by PhD Dr. Emanuel and Dr. Battin in the New England Journal of Medicine, the potential financial implications of PAS show that, although the reality of the cost reductions as a result of a diminished number of people using life support is not extremely significant, they can still create incentives for insurers to subtly promote physician-assisted suicide to patients facing long and expensive treatment. Specifically, legalizing PAS may lead to vulnerable patients being taken advantage of and pressured by healthcare systems wanting to reduce fiscal costs (Emanuel and Battin). The relevance of this is disputed by Emanuel, who affirms that PAS, in countries where its use is less regulated, is not commonly employed, thus contesting the previously noted concern regarding insurance companies.

Legislation in the US

The legality of physician-assisted suicide in the United States reflects a system of decentralization, mainly shaped by the precedent of the Supreme Court. In Vacco v. Quill (1997) and Washington v. Glucksberg (1997), the Supreme Court ruled that there is no constitutional right to PAS. As a result, the legality of PAS varies state by state (Cona). Oregon was the first to implement a "Death with Dignity" law in 1997, following the rulings noted above, and has served as the model for legislation elsewhere, with the main consistency being that self-administration is required (Oregon Health Authority). As of 2026, PAS is legal in California, Colorado, Delaware, the District of Columbia, Hawai'i, Illinois, Maine, Montana, New Jersey, New Mexico, Oregon, Vermont, and Washington. Additional states like Indiana, Massachusetts, Minnesota, Missouri, New Hampshire, New York, North Carolina, Pennsylvania, Rhode Island, Tennessee, and Virginia are currently considering similar legislation, though Illinois and the District of Columbia are considering a return to prohibition (Death with Dignity). This legislative fluidity in medicine mirrors the findings of Daikos in his highlighting of the transformed interpretations of the Hippocratic Oath throughout the years, as both reflect a field that is perpetually negotiating inherited ethics and contemporary medical realities. Across all states, it is clear that debates regarding autonomy, medical ethics, vulnerability, and regulations are always ongoing and shifting, with no law being truly set in stone when it comes to end-of-life policy in the United States.

Gap

While there are existing studies on ethical justifications, objections, and legislation surrounding physician-assisted suicide, several gaps remain understudied. Many sources focus on international cases without examining how ethics apply to public opinion and individual justification. There is also a lack of awareness of the views of medical professionals, who are arguably the most directly involved group. Furthermore, current literature does not contextualize data on perspectives within established ethical theories, and does not compare the differing laws of countries. As a result, there is little understanding of whether the attitudes of medical professionals in one concentrated area, Los Angeles County, are reflected in ethical frameworks and policies. This study aims to fill this gap by examining the extent to which the views of medical practitioners align with biomedical ethics and PAS legislation.

Methodology

Hypothesis and Introduction

After understanding the aforementioned context, the researcher found two primary hypotheses:

1. Medical professionals with greater clinical experience will demonstrate increased openness toward PAS compared to those with less experience.
2. A majority of those surveyed will express support for the development of formal legal and ethical frameworks that clearly define the boundaries of end-of-life care.

The first hypothesis is based on the premise that professionals with more extensive experience in the medical field have had greater exposure to terminal illnesses and patient suffering. Through repeated interactions with patients facing irreversible conditions, these individuals may develop a deeper understanding of autonomy, dignity, and quality of life. As a result, they may be more inclined to view PAS as a compassionate option.

The second hypothesis comes from the obvious growing global discourse around end-of-life care and the inconsistencies with which countries have addressed PAS. Given the firsthand experience with ambiguities and ethical dilemmas that arise in clinical settings, many medical professionals may recognize the need for clearer, standardized policies. The absence of well-defined legal frameworks creates uncertainty for both patients and practitioners. Therefore, respondents will likely favor the establishment of procedural standards for PAS in all countries, as a means of ensuring ethical accountability as well as patient protection.

This study examines the perspectives of LA County medical professionals regarding physician-assisted suicide, and how they align with ethical frameworks as well as the policies of a select range of countries. The approach is exploratory, which is the best fit for investigating patterns and beliefs that can contribute to future findings. The researcher used a non-experimental and mixed-method design. Qualitative factors, such as thematic analysis of open-ended survey questions or published articles, allowed the researcher to explore the reasoning behind the respondents' views. Quantitative measures, like Likert-scale questions in surveys or frequency counts in existing studies, complemented this by showing general levels of acceptance or opposition.

Population

The study involves medical professionals practicing in LA County who voluntarily completed a 15-20 minute survey. This ensured a comprehensive overview of varying perspectives of practitioners and how their field and experience affected their views. Unlike prior research, this allows for many opinions in different specializations to be taken into account and to be able to cite professional logic. Experts were identified through professional academic networks, affiliations, and student reporting. Additionally, the study incorporated primary and secondary source analysis.

Instruments

The survey questions were split up into five sections in addition to informed consent (Appendix A), with questions inspired by a multitude of preexisting studies that were credited.

1. Demographics: This section asks general questions about your professional role and clinical experience. These questions help provide context for interpreting responses across different medical backgrounds.
2. General Views (Likert-Scale): This section includes statements about physician-assisted suicide. You will be asked to indicate your level of agreement using a Likert scale.

3. Ethical Policies: This section explores how different ethical principles and frameworks apply to physician-assisted suicide. Responses will help identify which ethical considerations most strongly influence professional views.
4. Vignettes: This section presents hypothetical clinical scenarios involving end-of-life decision-making. These vignettes are not intended to represent actual patients but are used to understand how ethical judgments may vary across different clinical contexts.
5. Reflection: This section provides an opportunity to elaborate on your views or raise considerations that may not have been addressed earlier. Responses would help contextualize prior answers.

Demographics		
Question	Answer Type/Options	Source
What is your primary professional role?	Multiple Choice ● Physician (MD/DO) ● Nurse Practitioner ● Physician Assistant ● Registered Nurse ● Other (Free Response)	Self-defined
What is your primary area of practice?	Multiple Choice ● Primary Care ● Oncology ● Palliative/Hospice Care ● Emergency Medicine ● Intensive Care ● Surgery ● Other (Free Response)	Self-defined
How many years have you been in clinical practice?	Multiple Choice ● 0-5 ● 6-10 ● 11-20 ● 21+	Self-defined
How often do you care for patients with terminal or life-limiting illnesses?	Likert-Scale ● Very often ● Often ● Sometimes ● Rarely ● Never	Self-defined
How familiar are you with physician-assisted suicide laws?	Likert-Scale ● Very familiar ● Somewhat familiar ● Slightly familiar ● Not familiar	Self-defined

Table 1: Demographic section of the survey.

General Views (Likert Scale)

Question	Answer Type/Options	Source
Physician-assisted suicide (where the patient self-administers medication) is ethically different from euthanasia (where the clinician administers the medication).	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Definition by Gesang
Physician-assisted suicide is ethically acceptable in some circumstances.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Self-defined
The Hippocratic Oath should be reinterpreted in a modern-day context, thus allowing for flexibility in end-of-life care.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Indla and Radhika
Respect for patient autonomy supports allowing physician-assisted suicide.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Kittay
The principle of "do no harm" conflicts with physician-assisted suicide.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Indla and Radhika
I clearly distinguish between euthanasia, physician-assisted suicide, and withdrawal of life-sustaining treatment.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Daikos
Do Not Resuscitate (DNR) orders are ethically distinct from assisted dying.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Daikos
Explain any response if you wish to elaborate.	Free Response	

Table 2: General views section of the survey.

Ethical Policies		
Question	Answer Type/Options	Source
Which ethical framework most closely aligns with your personal view on physician-assisted suicide?	Multiple Choice ● Principlism: based on the autonomy, beneficence, nonmaleficence, and justice of and for a patient. ● Utilitarianism: the best action is the one that creates the most happiness or good for the most people, while minimizing suffering for everyone involved ● Kantian/Deontological ethics: the morality of an action comes from whether it follows a universal moral rule, not its results. Intention over outcome. ● Virtue ethics: an action is only right if it is an action that a virtuous person would carry out in the same circumstances ● Sanctity of life: the belief that human life is inherently sacred, holy, and possesses infinite value. ● Unsure ● Other (Free Response)	Definitions by Baranzke; Bulger; Hursthouse and Pettigrove; Kranak; Singer
Of the four principles of biomedical ethics, which do you believe is most important in end-of-life decisions?	Multiple Choice ● Autonomy: respecting the patient's right to make decisions about their own body and life. ● Beneficence: acting to relieve suffering and promote well-being. ● Nonmaleficence: avoiding causing harm. ● Justice: ensuring fair access and treatment in end-of-life care.	Definitions by Bulger
Please explain your responses.	Free Response	Self-defined
Patients with decisional capacity should have ultimate authority over all end-of-life decisions.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Bulger
Vulnerable patients may feel pressured to choose assisted dying.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Emanuel and Battin
Medical professionals should have the right to object if a patient legally chooses physician-assisted suicide.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Self-defined

All countries should have a framework outlining the acceptability of physician-assisted suicide.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Inspired by Asai et al.
Explain any response if you wish to elaborate.	Free Response	Self-defined

Table 3: Ethical policies section of the survey.

Vignettes		
Question	Answer Type/Options	Source
A 68-year-old patient has terminal cancer. They are in severe pain that cannot be controlled, despite palliative care. They ask their caretaker to prescribe them a medication that will let them end their life peacefully at home. This is physician-assisted suicide, not euthanasia.		Self-defined
In this case, physician-assisted suicide is ethically acceptable.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Self-defined
Why?	Free Response	Self-defined
Would your answer have changed if the patient had asked for euthanasia instead of physician-assisted suicide?	Multiple Choice ● Yes ● No	Inspired by Gesang
A 35-year-old has a long-term, non-terminal condition that causes constant pain and makes daily life very difficult. There are some pills they can take, but they refuse to be reliant on them. They are not dying, but feel their quality of life is unbearable. They request assisted suicide.		Self-defined
In this case, physician-assisted suicide is ethically acceptable.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Self-defined
Why?	Free Response	Self-defined
Would your answer have changed if the patient had asked for euthanasia instead of physician-assisted suicide?	Multiple Choice ● Yes ● No	Inspired by Gesang

A 45-year-old patient has an irreversible neurological condition causing total paralysis and loss of independence. It is not terminal. The patient is mentally competent and requests assisted dying.		Self-defined
In this case, physician-assisted suicide is ethically acceptable.	Likert-Scale ● Strongly Agree ● Agree ● Neutral ● Disagree ● Strongly Disagree	Self-defined
Why?	Free Response	Self-defined
Would your answer have changed if the patient had asked for euthanasia instead of physician-assisted suicide?	Multiple Choice ● Yes ● No	Inspired by Gesang
After having read the above vignettes, would any of your answers have changed if physician-assisted suicide were considered completely legal, though not necessarily ethical, for every case?	Multiple Choice ● Yes ● No	Self-defined
Why?	Free Response	Self-defined
Explain any response if you wish to elaborate.	Free Response	Self-defined

Table 4: Vignette section of the survey.

Reflection		
Question	Answer Type/Options	Source
Do you believe that, consciously or subconsciously, your religious beliefs influenced any of your responses? Why or why not?	Free Response	Inspired by Sinha et al.
What ethical concern most influences your view on physician-assisted suicide? Why?	Free Response	Self-defined
Is there anything you feel is often misunderstood in public debates about assisted dying?	Free Response	Self-defined

Table 5: Reflection section of the survey.

Implementation

The researcher was able to implement this study through two primary methods. The first was through distributing the survey to hospitals via professional and academic networks. The second was by requesting referrals through the researcher's community and school to identify additional LA County practitioners willing to participate. Respondents were given a time window of three weeks to fill out the questionnaire before it closed.

Data Analysis Method

The document analysis involved reviewing governmental frameworks, ethical policies, and historical commentaries to contextualize survey responses. Each philosophical framework was selected based on its relevance to the medical field, according to medical journals. The researcher also selected nine countries, spanning each inhabited continent, with additional weight given to nations that had made substantive legislative progress in end-of-life care. They then compared the legislative framework of each country to the aforementioned philosophies to determine, by proxy, the governing positions that each medical professional most agreed with.

Ethical Policy	Definition
Utilitarianism	Actions are judged by their consequences to minimize suffering. Often focuses on relief and quality of life as key ethical justifications. (Singer)
Kantian Deontology	Emphasizes duty and moral law over consequences. Intentionally ending a life violates the duty of preserving life (Kranak)
Principlism	Four main principles of autonomy, beneficence, nonmaleficence, and justice. Key philosophy of biomedical ethics. (Bulger)
Virtue Ethics	Main values are compassion, wisdom, and integrity. Importance of eudaimonia, which is human flourishing, the role of moral exemplars, and the balance of practical wisdom, known as phronesis. (Hursthouse and Pettigrove)
Sanctity of Life Ethics	Human life has inherent value and should never be intentionally ended. Grounded in religious or moral views. (Baranzke)
Ethics of Care	Reconceptualization of care and dependency so that systems of justice respect dignity and not just independence. Supports decisions aimed at alleviating suffering while honoring the dignity and needs of the individual. (Kittay)

Table 6: Selected ethical frameworks and definitions.

Country	Framework	Closest Ethical Policy
Canada (Government of Canada)	Under the Medical Assistance in Dying Bill C-14: Active Euthanasia - Legal PAS - Legal Passive Euthanasia - Legal	Principlism

France (De Nonneville et al.)	Under Claeys-Leonetti Law: Active Euthanasia - Illegal PAS - Illegal Passive Euthanasia - Legal *Important to note that it is currently under revision.*	Ethics of Care
Netherlands (Government of the Netherlands)	Under the Termination of Life on Request and Assisted Suicide Act: Active Euthanasia - Legal PAS - Legal Passive Euthanasia - Legal	Utilitarianism
Belgium (Government of Belgium)	Under the Belgian Act on Euthanasia of 2002: Active Euthanasia - Legal PAS - Not specifically regulated, very rare Passive Euthanasia - Legal	Utilitarianism
Colombia (López Benavides)	Under jurisprudence ruling C-239: Active Euthanasia - Legal PAS - Illegal Passive Euthanasia - Legal	Virtue Ethics
Australia (ELDAC Editors)	Under the Voluntary Assisted Dying Acts: Active Euthanasia - Illegal PAS - Legal (In selected states) Passive Euthanasia - Legal	Principlism
South Korea (Yun et al.)	Under the Life-Sustaining Treatment Decision Act: Active Euthanasia - Illegal PAS - Illegal Passive Euthanasia - Legal	Ethics of Care
India (Supreme Court of India)	Under the Common Cause v. Union of India 2018 case ruling: Active Euthanasia - Illegal PAS - Illegal Passive Euthanasia - Legal with strict precautions	Sanctity of Life
South Africa (Kotzé and Roos)	Under the National Health Act 61 of 2003: Active Euthanasia - Illegal PAS - Illegal Passive Euthanasia - Legal with strict precautions	Kantian Deontology

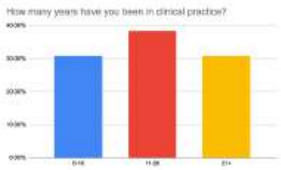
Table 7: Selected country policies and aligning ethical frameworks.

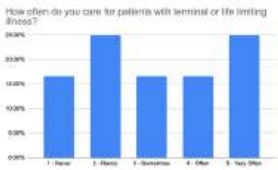
Findings & Analysis

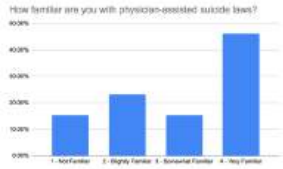
As previously discussed, different methods were used to analyze the data collected from the LA County medical professional survey. To understand the pattern of responses in a comprehensive context, descriptive statistics were used to examine the Likert-scale items, while the open-ended questions were reviewed for repeated responses and broad consensus. In total, the sample of responses included respondents from diverse healthcare backgrounds, including

nurses, physicians, speech and language pathologists, psychologists, and dentists, all with varying years in clinical practice. This variation in specialization also produced notable differences in the frequency with which respondents encountered terminally ill patients, a factor that proved relevant to several patterns in the data. Some graphs are more significant in helping to answer the overall research inquiry and will therefore receive extended analysis.

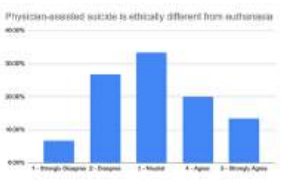
Demographics:

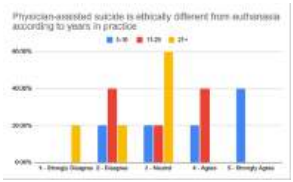




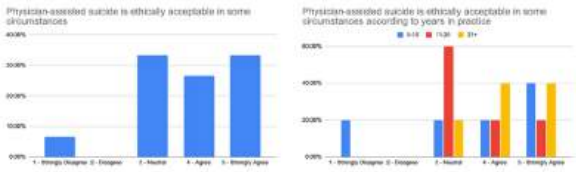


General Views:

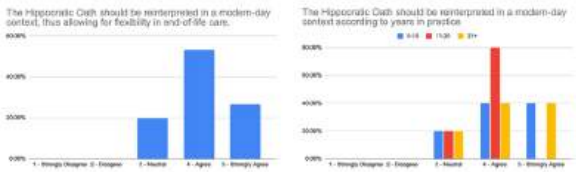




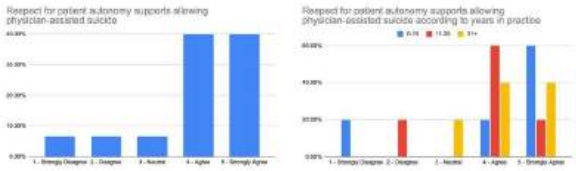
As seen in the histogram above, 100% of surveyed medical professionals working for 21+ years leaned towards the middle to the disagreement side of the scale. In comparison, this was true for 60% of those with 11-20 years and only 40% for 0-10. This pattern suggests that clinical experience shapes how strongly professionals distinguish between PAS and euthanasia, a distinction that lessens with time. Respondents with 21+ years showed the greatest concentration of responses on the disagreement end of the Likert scale, reflecting the reality that prolonged exposure to terminal patient care collapses theoretical distinctions under the weight of clinical practice. As both interventions result in death, experienced practitioners may be less persuaded that the patient's role in administering the medication constitutes a morally decisive difference. This aligns with Gesang's framework, in which the distinction between PAS and active or passive euthanasia is already tenuous. The data suggest that clinical exposure reinforces skepticism.

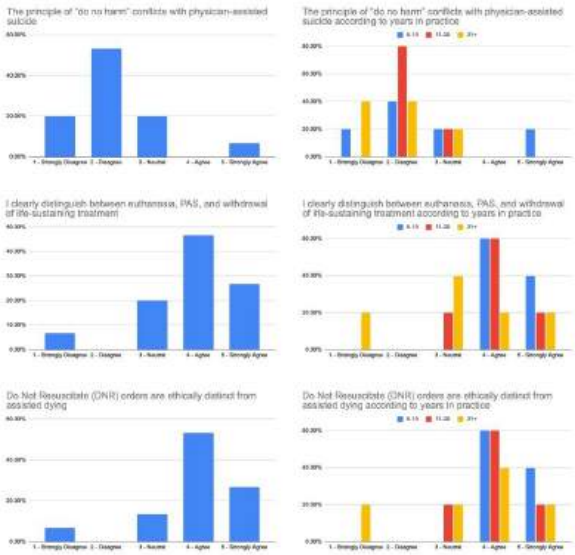


The data show that almost all respondents leaned towards the middle to agreement side of the scale for acceptability, with the exception of some outliers who stated that PAS "should never be presented as an acceptable solution to suffering" (Respondent 3). The notable hesitancy among those with 11-20 years of practice, which represents the middle cohort among those surveyed, may reflect a transitional phase in professional identity: the practitioner has enough experience to have encountered the complexities of the situation, but not enough sustained exposure to have a solid conviction. Those at the extremities of tenure, by contrast, exhibited more definitive views, which is consistent with the hypothesis that clinical depth clarifies moral judgment on this issue. Those with greater experience in this sample seem to have more firmly established opinions about the permissibility of PAS in relation to ethics. In short, the trend shifts toward agreement among respondents with more extensive clinical backgrounds.

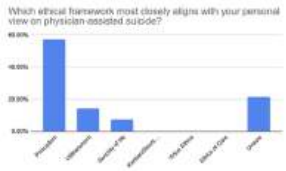


This graph indicates that there is a clear consensus that the Hippocratic Oath should be reinterpreted in a modern-day context, regardless of years in practice. The 11-20 years group appears to be the most supportive, with 80% outright agreeing, while the 0-10 and 21+ years groups are evenly split between the latter three categories on the Likert scale. There were no responses disagreeing, indicating that medical professionals in this sample generally recognize the need for classic texts to be updated. This is extremely significant as it shows that the Oath's inadequacy for contemporary end-of-life care is not a generational or experimental position, but a professional conclusion. This finding reinforces the second hypothesis of the research, and also echoes the verdicts of both Daikos and Indla and Radhika, who each argued that the Oath was never designed to accommodate the current ethical terrain, as previously mentioned.

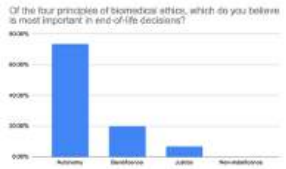




Ethical Policies:

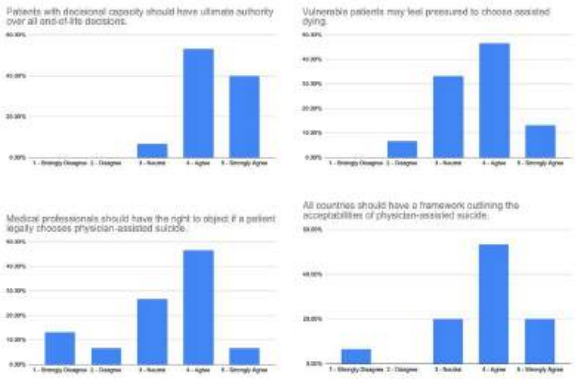


According to the graph above, principlism seems to be the most popular framework, though utilitarianism and sanctity of life each claimed 14.29% and 7.14% of responses, respectively. As previously explained, principlism is an ethical framework that leans towards patient self-determination in debates regarding topics such as euthanasia. Its dominance is consistent with its flexible design; rather than having a fixed hierarchy, it provides four key principles (autonomy, beneficence, nonmaleficence, and justice) that clinicians can weigh depending on the context. In a field defined by individual patient circumstances, this malleability is an advantage. The fact that it was a frontrunner against more rigid frameworks, like Kantian deontology (0%) and sanctity of life (7.14%), suggests that LA County practitioners are inclined to select ethical reasons that can accommodate the complexity of their professions.

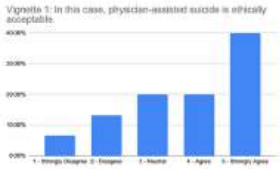


This question was included based on the fact that these pillars are often seen as the core of biomedical ethics (Bulger). Autonomy was the most selected at 73.33%, showing the great emphasis that respondents placed on self-determination and the right to make decisions about personal care. Beneficence received 20% of the votes, and justice took 6.67%, a distribution that

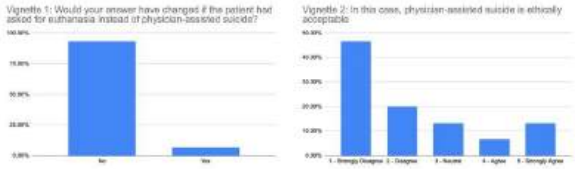
reveals something more specific than a general appreciation for autonomy. It suggests that, for this sample, the focus of ethical authority in end-of-life decisions lies with the patient rather than with the outcome or with systemic fairness. Notably, nonmaleficence received no responses as the primary principle, which is particularly striking given that it is central to Hippocratic tradition and to the most common objections to PAS.



Vignettes:

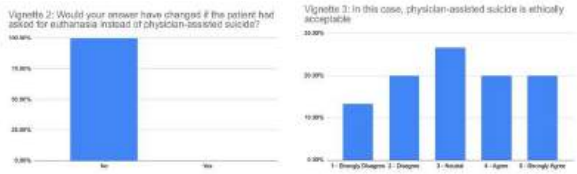


Respondents for this Vignette reached an overall agreement, approving PAS for the specific scenario. Although 20% leaned towards disapproval, stating that "PAS involves a deliberate act to end life, which conflicts with the fundamental ethical duty of physicians to preserve life and not intentionally cause death," 80% thought it was permissible (Respondent 3). The 80% approval rate aligns directly with the autonomy-first orientation identified in the ethical policies sections: when a patient is cognitively intact, in severe pain, and facing a terminal diagnosis, respondents overwhelmingly treated their wish to die as the determining factor for the course of action. The terminal condition appears to function as a threshold criterion; if it is not met, support collapses sharply, as seen in the upcoming Vignettes 2 and 3.

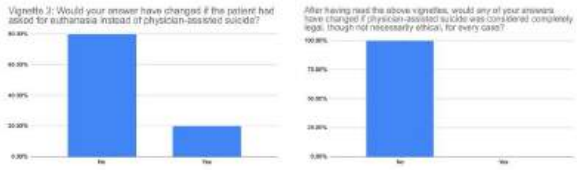


Respondents for this Vignette once again reached an overall agreement, disapproving of PAS for the specific scenario. In a reverse pattern, although 20% leaned towards approval, stating that "it is always the patient's wish" and highlighting the importance of autonomy, 80% considered it inappropriate for the situation (Respondent 7). The most common reasoning centered on the "non-terminal nature of the condition and the availability of alternative"

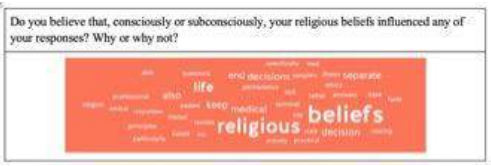
treatments (Respondent 4). This result reveals the limits of the autonomy principle as applied by this cohort: while 73.33% had selected autonomy as their primary ethical value, that commitment does not appear to be unconditional. When a non-terminal patient refuses available treatment and requests PAS, the majority of respondents declined to honor that request, thus suggesting that, in practice, beneficence and nonmaleficence act as implicit constraints on patient self-determination, even among those who nominally prioritize it.



Responses for this Vignette showed a different pattern than the previous one, with some discord among opinions. There is a lower standard deviation, indicating that most medical professionals leaned towards neutrality or taking less strong stances. One respondent's explanation reflects the common concern of multiple participants: “It is difficult to tease out a permanent state of mind vs. a period of depressed mood and grief after loss” (Respondent 5). This scenario occupied an ambiguous middle ground that neither the terminal case, Vignette 1, nor the non-terminal case, Vignette 2, could occupy. Paralysis that is irreversible but non-terminal tests the autonomy principle most severely: the patient has full decisional capacity and no prospect of recovery, yet the condition does not meet the aforementioned threshold that motivated approval in Vignette 1. The convergence toward neutrality suggests that the ethical frameworks respondents had originally most identified with are not as clearly applied in this scenario.



Reflection:



The answer to this question was always “no,” suggesting that participants generally saw their views on PAS as grounded in unbiased judgment, though this may be limited by self-reporting bias. The consistency of this response across all respondents is notable, but it must be interpreted with caution. Self-reporting bias is particularly present on questions involving religion, where social desirability pressures are often strong. That said, if taken at face value, the unanimous rejection of religious influence has meaningful policy implications. It essentially

would mean that it is possible to separate science from faith, thus supporting the feasibility of developing PAS legislation that does not require consensus on a theological level.

The answers to this question were varied, though decisional capacity was a prominent one. Specifically, many medical professionals expressed concern that patients requesting PAS would not be in the correct state of mind to do so, which points directly to the tension between autonomy and nonmaleficence identified throughout this study. Overall, the responses were generally divided between protecting patients from harm, with some directly discussing nonmaleficence and safeguards, while others focused on quality of life. The diversity of reasoning is itself a finding; even among respondents who agreed that PAS is ethically acceptable in some circumstances, the ethical concern they named as most influential varied considerably. This proposes that support for PAS, at least in this sample, does not reflect a unified ethical position, but rather a convergence of different principles arriving at similar conclusions through distinct paths.

Responses to this question often highlight the misunderstandings that are present in the realities of end-of-life care, patient experience, and ethical nuances. Several surveyed medical professionals pointed out the underacknowledged “impact of terminal illness on patients and their families,” which debates often ignore (Respondent 13). Another explained that “many pain medications shorten life at the same time as they provide pain relief,” raising a point that the existing literature has largely neglected: palliative treatments already produce outcomes that are functionally similar to PAS through side effects, yet face far less ethical scrutiny (Respondent 9). One response is particularly notable: “Bodies often outlive brains. Death can be a blessing” (Respondent 10). This formulation highlights what public debates consistently fail to account for. For some patients, the central point of suffering is neurological and existential instead of physical, and the standard medical goal of preserving life may, in those cases, be operating against patient welfare.

Conclusion

This study originally set out to compare the views of LA County medical professionals regarding physician-assisted suicide to establish ethical policies and legal frameworks globally.

After researching and analyzing survey responses, the general conclusion is that most medical professionals surveyed consider PAS ethically acceptable under narrowly defined conditions, specifically when a patient is terminally ill, cognitively intact, and has exhausted all other available alternatives. This was especially true for professionals who had been in clinical practice, which is likely due to the fact that they have experienced more patient interactions, thus having a greater understanding of the scope of the topic and how people are affected. This reinforces the researcher's original hypothesis and relates back to Kittay and Emanuel, who both argued for PAS based on listening to and respecting patient decisions. The preceding survey data, which include Likert-scale responses, vignette scenarios, and open-ended reflections, revealed a clear preference for principlism as an ethical foundation for medical beliefs over others, such as utilitarianism, sanctity of life, Kantian deontology, and virtue ethics. Within that framework, autonomy emerged as the dominant pillar, selected by 73.33% of respondents as the principle most influential in end-of-life decisions. The contrast in support for varying vignettes fluctuated between terminal and nonterminal cases in a pattern hypothesized by the researcher and reflects the conflict between patient self-determination and nonmaleficence, which is stated as a professional duty in the Hippocratic Oath. However, the unanimous consensus across all experience cohorts that the Oath requires reinterpretation suggests that this conflict is widely recognized and that practitioners are relatively non-conforming to letting a 5th-century B.C. text dictate the boundaries of 21st-century clinical ethics (David Geffen School of Medicine). Moreover, the unanimous rejection of religious influence highlights how medical professionals are able to separate their personal faith when confronted with nuanced topics, which would mean that laws in the future can be created secularly as well. When comparing the ethical beliefs of those surveyed to existing country policies, it can be determined that the openness of most aligns most closely with countries like Canada, the Netherlands, and parts of Australia, refuting the restrictive policies of South Korea, South Africa, and India. To address ongoing end-of-life debates, medical education and policy must integrate these professional perspectives.

Implications

The findings suggest that medical training and policy discussion play a critical role in shaping the beliefs of medical professionals regarding PAS. It is crucial to place more emphasis on the definitions of varying types of terminal care, including PAS, active and passive euthanasia, palliative care, and Do Not Resuscitate orders, to better prepare clinicians to navigate ethical grey areas without defaulting to absolutism. As stated earlier, the support for PAS declined in non-terminal cases, implying that legislation modeled on the Oregon framework, which requires terminal diagnosis, may reflect something closer to the ethical intuitions of practicing clinicians (Oregon Health Authority). Policymakers drafting end-of-life legislation could use this alignment as evidence that such criteria reflect the moral reasoning of those who would be expected to help carry out PAS. Many countries have yet to enter this debate, but the results of this study have shown that nations need to have defined policies to limit ambiguity and uncertainty about what is acceptable. In that process, medical professionals themselves can

provide valuable insight, as their experience places them in a strong position to make educated decisions.

Limitations

While this research offers much meaningful insight into the topic, some limitations must be acknowledged. The sample size that the survey was able to reach does not encompass all medical professionals in LA County, which means that findings should be treated as exploratory rather than completely representative. They illuminate patterns worth investigating on a grander scale, but cannot be taken as descriptive of the broader LA County medical community. Additionally, survey responses were self-reported and may be subject to social desirability bias, especially on sensitive topics like ethical duties and religion. The latter is especially susceptible, since implicit religious influences on moral reasoning can operate without one being aware, and respondents have no reliable way to assess the extent to which their views have been shaped by beliefs that they report to be unimportant when it comes to their professional judgement.

Future Research

Building on the findings of this study, future research should examine a broader range of medical professionals across multiple states or countries from more diverse specialties to assess various regional attitudes towards PAS. In addition, they should narrow their focus on significant factors influencing end-of-life decision-making, such as the role of different methods of training in affecting the formation of beliefs about PAS. This can be done through the use of a similarly structured survey as the one used in this study to be able to compare the full spectrum of views. Additionally, follow-up interviews with a larger subset of respondents could also provide deeper insights into the reasoning behind ethical stances. One question raised is whether the ethical intuitions identified here are stable across time. Essentially, whether the same practitioners would respond differently after additional years of exposure to terminal cases, or after practicing under a legalized PAS framework. A longitudinal design would allow researchers to track how professional ethics evolve in response to both clinical experience and changing legal environments. Ultimately, this would ensure the goal of compassionate end-of-life care for future generations.

Works Cited

Asai, A., Okita, T., Shimakura, Y., Tanaka, M., & Fukuyama, M. (2023). Japan Should Initiate the Discussion on Voluntary Assisted Dying Legislation Now. BMC Medical Ethics, 24(1). https://doi.org/10.1186/s12910-023-00886-0

Baranzke, H. (2012). "Sanctity-of-Life"—A Bioethical Principle for a Right to Life? Ethical Theory and Moral Practice, 15(3), 295–308. https://doi.org/10.1007/s10677-012-9369-0

Bulger, J. W. (2007). Principlism. Teaching Ethics, 8(1), 81–100. https://doi.org/10.5840/tej2007816

Cona, N. (2022). Physician-Assisted Suicide Legalization in the United States. Sacred Heart University. https://digitalcommons.sacredheart.edu/cgi/viewcontent.cgi?article=1808&context=acadfest

Daikos, G. K. (2007). History of Medicine: Our Hippocratic Heritage. International Journal of Antimicrobial Agents, 29(6), 617–620. https://doi.org/10.1016/j.ijantimicag.2007.01.008

David Geffen School of Medicine. (2023, July 20). Modern Hippocratic Oath: UCLA Med School. UCLA Medical School. https://medschool.ucla.edu/blog-post/modern-hippocratic-oath-holds-the-underlying-values-of

Death with Dignity. (2026). Death with Dignity: U.S. Legislative Status State Map. https://deathwithdignity.org/states/

De Nonneville, A., Marin, A., Chabal, T., Tuzzolino, V., Fichaux, M., & Salas, S. (2016). End-of-life practices in France under the Claeys-Leonetti Law: Report of three cases in the oncology unit. Case Reports in Oncology, 9(3), 650–654. https://doi.org/10.1159/000450940

ELDAC Editors. (2024, August 28). End of Life Law - Overview of Voluntary Assisted Dying. ELDAC. https://www.eldac.com.au/Our-Toolkits/End-of-Life-Law/Voluntary-Assisted-Dying/Overview

Emanuel, E. J. (2002). Euthanasia and Physician-Assisted Suicide. Archives of Internal Medicine, 162(2), 142. https://doi.org/10.1001/archinte.162.2.142

Emanuel, E. J., & Battin, M. P. (1998). What are the Potential Cost Savings from Legalizing Physician-Assisted Suicide? New England Journal of Medicine, 339(3), 167–172. https://doi.org/10.1056/nejm199807163390306

Gesang, B. (2007). Passive and Active Euthanasia: What is the Difference? Medicine, Health Care and Philosophy, 11(2), 175–180. https://doi.org/10.1007/s11019-007-9087-x

Government of Belgium. (2002). The Belgian Act on Euthanasia of May 28th 2002. Ethical Perspectives, 9(2), 182–188. https://doi.org/10.2143/ep.9.2.503856

Government of Canada, D. of J. (2023, February 2). Legislative Background: Medical Assistance in Dying (Bill C-14, as assented to on June 17, 2016). Part 2 – Consideration of International Medical Assistance in Dying Regimes. https://www.justice.gc.ca/eng/rp-pr/other-autre/adra-amsr/p3.html

Government of the Netherlands, M. of Health, Welfare, and Sport. (2025, April 29). Is Euthanasia Legal in the Netherlands?. Euthanasia | Government.nl. https://www.government.nl/topics/euthanasia/is-euthanasia-allowed

Hursthouse, R., & Pettigrove, G. (2022, October 11). Virtue Ethics. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/ethics-virtue/?ref=grokk.ist

Indla, V., & Radhika, M. (2019). Hippocratic Oath: Losing Relevance in Today's World? Indian Journal of Psychiatry, 61(10), 773. https://doi.org/10.4103/psychiatry.indianjpsychiatry_140_19

Kittay, E. F. (2011). The Ethics of Care, Dependence, and Disability*. Ratio Juris, 24(1), 49–58. https://doi.org/10.1111/j.1467-9337.2010.00473.x

Kotzé, C., & Roos, J. L. (2022). End-of-Life Care in South Africa: Important Legal Developments. South African Journal of Psychiatry, 28. https://doi.org/10.4102/sajpsychiatry.v28i0.1748

Kranak, J. (2019, December 9). Kantian Deontology. Introduction to Philosophy: Ethics. https://spscc.pressbooks.pub/ucfcopy/chapter/kantian-deontology/

López Benavides, L. L. (2018). The Right to Die with Dignity in Colombia. Forensic Research; Criminology International Journal, 6(6), 426–429. https://doi.org/10.15406/frcij.2018.06.00239

Nguyen, D. (2016). Brain Death and True Patient Care. The Linacre Quarterly, 83(3), 258–282. https://doi.org/10.1080/00243639.2016.1188472

Oregon Health Authority. (n.d.). Oregon's Death with Dignity Act. Oregon Health Authority : Oregon's Death with Dignity Act : Death with Dignity Act : State of Oregon. https://www.oregon.gov/oha/PH/ProviderPartnerResources/EvaluationResearch/DeathwithDignityAct/Pages/index.aspx

Rurup, M. L., Smets, T., Cohen, J., Bilsen, J., Onwuteaka-Philipsen, B. D., & Deliens, L. (2011). The First Five Years of Euthanasia Legislation in Belgium and the Netherlands: Description and Comparison of Cases. Palliative Medicine, 26(1), 43–49. https://doi.org/10.1177/0269216311413836

Singer, P. (2003). Voluntary Euthanasia: A Utilitarian Perspective. Bioethics, 17(5–6), 526–541. https://doi.org/10.1111/1467-8519.00366

Sinha, V., Basu, S., & Sarkhel, S. (2012). Euthanasia: An Indian Perspective. Indian Journal of Psychiatry, 54(2), 177. https://doi.org/10.4103/0019-5545.99537

Spurdle, M., Henderson, P., Worboys, S., Scrimgeour, T., & Spiekermann, D. (2021, November 1). What is the Difference Between Euthanasia and Assisted Suicide?. Maxim Institute. https://www.maxim.org.nz/article/faq1-euthanasia-and-assisted-suicide/

Supreme Court of India. (2023, December 26). Euthanasia and the Right to Die with Dignity. Supreme Court Observer. https://www.scobserver.in/cases/common-cause-euthanasia-and-the-right-to-die-with-dignity-case-background/

World Health Organization. (n.d.). History of Polio Vaccination. World Health Organization. https://www.who.int/news-room/spotlight/history-of-vaccination/history-of-polio-vaccination

Yun, Y. H., Sim, J.-A., Choi, Y., & Yoon, H. (2022). Attitudes Toward the Legalization of Euthanasia or Physician-Assisted Suicide in South Korea: A Cross-Sectional Survey. International Journal of Environmental Research and Public Health, 19(9), 5183. https://doi.org/10.3390/ijerph19095183

Plural Emotion Readings in Short Social Video Clips: A Descriptive Study

By Evelyn Kim

Abstract

Emotion recognition datasets used to train affective computing systems typically assign each stimulus a single "correct" emotion. This convention rests on a testable assumption: that when two thoughtful readers watch the same social moment and arrive at different emotion labels, one of them is wrong. In a preliminary study with 43 adult participants across seven countries, 442 non-excluded responses were collected on 33 short social video clips (10–30 seconds each). Participants selected among 14 emotions or a "mixed" option to indicate multiple simultaneous emotions, attributed each reading to specific perceptual cues, and rated their confidence. The observations point to two levels of plurality in the human response data. **Across readers:** on the median clip, the most-picked emotion was chosen by 70% of readers (IQR: 50–76%), with roughly five distinct emotions mentioned per clip; eight of 33 clips (24%) reached the ≥ 80% agreement level used here to define a consensus tier. **Within individual responses:** 24% of responses used the "mixed" option to select more than one simultaneous emotion, with a mean of 1.42 emotions selected per response, and mixed responses carried nearly the same confidence rating as single-emotion responses (4.24 vs 4.44 on a 5-point scale). Breaking mixed responses down by confidence level, the mixed rate did fall as confidence rose (from 37% at confidence 3 to 20% at confidence 5), but roughly one response in five at the highest confidence level was still mixed, so multi-emotion selection is not entirely accounted for by reader uncertainty. Contrary to a naive intuition, clips with heavier verbal content did *not* show higher inter-reader agreement; the descriptive pattern pointed the other way. Together the observations indicate that the human response distribution on this corpus is not well represented by a single canonical label, and that single-label representation under-specifies the observed human response distribution on this corpus. The response-level dataset (adult participants) is released to enable replication and reuse.

1. Introduction

1.1 The single-label convention and its blind spot

Emotion recognition systems have quietly become part of the everyday infrastructure of contemporary life. They monitor student engagement in classrooms, feed affective cues into autism-support wearables, help driver-alertness pipelines detect fatigue, and increasingly mediate the emotional register of conversational agents. Each of these deployments rests on a training pipeline with a common architectural choice: for every video, image, or vocal clip in the training set, a single "correct" emotion label is attached, and the model learns to reproduce that label. AffectNet (Mollahosseini et al.), the largest publicly-available facial-expression corpus, provides one categorical label per face (alongside continuous valence/arousal annotations), with the categorical label typically used as the top-1 target in supervised training; RAVDESS and

Affectiva's face database (Zijderveld) follow the same categorical-label convention; contemporary reviews of the affective-computing pipeline (Wang et al.; Latif et al.) treat single-label supervision as the default methodological premise.

The convention has practical roots. Single-label supervision is straightforward to compute against, integrates cleanly with the accuracy-focused benchmarks the field has built up over decades, and permits large-scale automated evaluation. Inter-rater disagreement in early emotion corpora was routinely filtered out through reliability thresholds, through majority voting, or through expert re-annotation designed to converge on a single canonical label per item (Landowska et al.). These conventions made engineering progress tractable and, over time, hardened into methodological assumptions about the phenomenon itself.

Those assumptions are worth stating explicitly. The single-label convention assumes that when two thoughtful readers watch the same twenty-second social interaction and arrive at different emotion labels, at least one of them must be wrong. This paper collects human readings on a small set of short social clips and asks, in the most direct descriptive terms, whether the human response distribution is well represented by that assumption.

1.2 Emotion perception as plural

A growing body of work in affective science challenges the single-label default and argues that the phenomenon under measurement admits multiple valid readings. Cowen and Keltner mapped twenty-eight distinct emotions conveyed by naturalistic facial expressions, a much larger emotion space than the classical six-basic set (Ekman) that most training corpora inherit, and their companion work on vocalization (Cowen et al.) documents a similar expansion of the space beyond canonical categories. Barrett et al. issued a formal challenge to the tradition of inferring emotion from human facial movements, concluding that facial configurations do not map reliably onto discrete emotional states. Barrett's broader constructivist program (*How Emotions Are Made*) reframes emotion perception as an active construction: a moment does not reveal its own emotional identity, and readers assemble one from context, prior experience, and the specific cues available to them at reading time.

If the constructivist view is right, then variance in reader labels is not measurement noise to be filtered away; it is the actual signal, an observable consequence of the fact that the construction process differs across readers. Whether the constructivist view is right is, in part, an empirical question, and one this paper takes up in exploratory descriptive form.

1.3 The empirical gap: from a literature review to a testable question

The present study is a direct empirical follow-up to a specific prior paper. Kim ("Artificial Intelligence"), reviewing AI-based emotional-intelligence training tools designed to support autistic learners, argued that the tools' effectiveness is structurally constrained by the single-label assumption of the datasets on which they are trained: if the training data enforces one emotion per moment while real social environments may support multiple plausible readings, a tool that succeeds on the benchmark may still fail in practice. That earlier paper closed with an open

question, whether the single-label assumption would survive direct scrutiny on how human observers actually read short social moments, and pointed forward to the empirical study that would test it.

The present paper is that study, reported here in its preliminary descriptive-exploratory form on the current sample. Between May and August 2026, the study collected 442 non-excluded responses from 43 adult participants across seven countries, each reading a subset of 33 short social video clips. The paper reports two families of descriptive observations on these 33 clips: **across-reader plurality** (how much do different readers converge on the same emotion for a given clip?) and **multi-emotion responses** (what fraction of individual responses explicitly report more than one simultaneous emotion?). Both patterns are informative about the single-label convention, and both can be described without formal statistical apparatus.

The study was conducted under a pre-registered protocol (Nosek et al.), committed publicly in advance of substantive data collection. The pre-registration specifies three hypotheses (H1 on plurality; H2 on verbal dependency; H3 on social complexity) at particular quantitative thresholds. This paper reports the descriptive observations that speak to H1 and H2 on the current adult sample; formal hypothesis tests, inter-rater reliability metrics (Krippendorff's α, Shannon entropy over the emotion vocabulary), and the H3 analysis (which requires more clips at the extremes than the current corpus contains) are reserved for a subsequent empirical paper on an expanded corpus that will supersede this preliminary report.

1.4 Contributions

This paper makes two contributions.

Empirical. The human response distribution on these 33 short social video clips is plural at two levels.

- **Across readers:** the modal emotion on the median clip is chosen by roughly seven readers in ten, and about five distinct emotions are mentioned per clip.
- **Within individual responses:** about 24% of responses explicitly report more than one simultaneous emotion, with confidence ratings nearly as high as those for single-emotion responses.

Both patterns point in the same direction and complicate the single-label convention.

Methodological. The response-level dataset (adult participants) is released, paired with cue attribution and confidence ratings for every response, under a permissive open license. The dataset preserves inter-rater variance rather than suppressing it: every reading is retained, every disagreement is observable, and no majority-vote reduction is performed at storage time.

2. Related Work

2.1 Emotion perception: categorical, dimensional, constructed

The classical categorical tradition in emotion research, developed most influentially by Ekman, posits a small set of basic emotions with cross-culturally recognizable facial expressions.

That tradition supplies the six- or seven-category label vocabulary that most emotion-recognition corpora inherit and has, for decades, offered the field a tractable common taxonomy. It is not, however, the only game in town. Russell's dimensional model treats emotion as continuous variation along valence and arousal axes rather than as membership in discrete categories. Cowen and Keltner recovered a substantially larger emotion space, twenty-eight distinct emotions conveyed by naturalistic facial expression, using data-driven approaches that do not presume a canonical taxonomy in advance.

Barrett's constructivist program (*How Emotions Are Made*; Barrett et al.) offers a further step: emotion perception is treated as an active construction assembled by the reader from context, prior experience, and cue availability, rather than as a passive reading of a fixed emotional signal on a face. Under this framing, both the number of relevant categories and their applicability to any given moment are variable. The present descriptive study is compatible with either the categorical or the constructivist reading of its data, and its observations speak to the practical downstream consequences of assuming single-label mapping.

2.2 Emotion-recognition datasets and single-label supervision

Large-scale emotion-recognition corpora (AffectNet, Mollahosseini et al.; Affectiva's face database, Zijderveld; and their peers) typically resolve inter-rater disagreement at annotation time, through some combination of majority voting, expert re-annotation, and reliability-based filtering. Reviews of the resulting affective-computing pipeline (Wang et al.; Latif et al.) generally present these aggregation practices as standard components of dataset construction; Landowska et al. review the same pipeline in the specific context of autism-support tools. The result is a training-data convention that assumes each item has a canonical answer, and a downstream evaluation convention (top-1 accuracy against a modal or expert-consensus label) that inherits the same assumption.

Contemporaneous work extends the same critique into the textual domain. Abbas and Scott collect human and machine emotion interpretations on short pieces of textual emotional content and document a pluralistic pattern: human annotators do not converge on a single label, and consensus-oriented detection systems inherit that mismatch. The present paper takes up the same descriptive question on a different stimulus modality (short social video clips rather than text) and adds an explicit multi-emotion response option so that within-response plurality is directly observable in addition to across-annotator plurality. The two studies are complementary reads of the same underlying pattern in two different media.

2.3 Autism, emotion perception, and technological interventions

Emotion-perception research in autistic populations spans facial-recognition assessments (Golan et al.), behavioral meta-analyses (Uljarevic and Hamilton), broader review syntheses of neurobehavioral evidence (Harms et al.), socially-assistive-robot interventions (Silvera-Tawil et al.), and meta-analyses of facial-emotion-training programs (Zhang et al.). Theoretical accounts,

most prominently Baron-Cohen's empathizing–systemizing framework, shape how these interventions are designed and evaluated.

The originating paper for the present study (Kim, "Artificial Intelligence") reviewed this literature and identified a common structural limitation: the AI-based tools evaluated consistently inherit the single-label assumption of their training data. The present study collects a general-adult-sample descriptive dataset that speaks to that assumption; clinical-sample extensions are deferred to future project phases.

2.4 Pre-registration in exploratory studies

Pre-registration commits a study's hypotheses, thresholds, and analysis pipeline to a public record before substantive data collection begins (Nosek et al.). The broader open-science and reproducibility movement (Munafò et al.) treats pre-registration as one component of a coordinated set of practices that includes open data, open code, and pre-analysis disclosure. The present paper follows a pre-registered protocol committed on the project's public pre-registration page in advance of substantive data collection.

3. Methods

3.1 Participants

At the data cutoff of 2026-08-14, 63 participants had registered on the platform. Of these, 50 had submitted at least one response (515 raw responses). One participant was flagged as an internal-test account and their responses were removed, leaving 49 participants and 504 responses. Of the 49 remaining participants, 43 were adults (aged 18 or older) and 6 were minors (ages 13–17). This paper reports on the 43 adult participants (442 non-excluded responses across 33 clips); the six minor participants were excluded before analysis because verified parental consent was unavailable (see Ethics Statement). All analyses below use this adults-only analytical sample.

Ages of the 43 adult participants ranged from 21 to 57 (M = 40.7, SD = 9.3). Country distribution: Singapore (n = 16), Republic of Korea (n = 14), United States (n = 6), Cambodia (n = 3), Vietnam (n = 2), India (n = 1), Hong Kong SAR (n = 1). Recruitment was through personal networks, an international-school community (Singapore American School), and public LinkedIn outreach.

Exclusion criteria. Internal-test profiles and responses individually flagged for response-timing or "select-all" cue patterns are excluded from all reported analyses.

Consent and eligibility. Registration on the study platform required participants to affirm that they were aged 13 or older and to consent to the study's data-use terms (Section 7). Only the adults-only subsample is analyzed and released in this report; verified parental consent was unavailable for the six minor registrants during this data-collection window.

IRB status. This is an independent citizen-science project conducted outside the jurisdiction of any Institutional Review Board (see Section 7 for the full ethics framework and applicability analysis).

3.2 Stimulus set

The stimulus set comprises 33 short social video clips selected from publicly-available film, television, and interview footage. Each clip was curated to a 10–30 second window centered on a single identifiable target person whose emotion the reader would be asked to interpret. Clips were assigned to three collection waves (11 clips each) and are analyzed jointly in this paper.

Each clip was coded by the curator on **verbal dependency** (low = 12 clips, medium = 15, high = 6): how much the emotional content depends on the linguistic channel relative to facial and paralinguistic channels. Each clip also carried a brief target-person description shown to readers ("a lady in red," "the coach on the left") to anchor whose emotion the reader was asked to interpret; the description is a spatial anchor only and does not carry emotion valence.

3.3 Emotion taxonomy

The response instrument uses a 14-emotion taxonomy plus one "mixed" response option:

happy_amused, proud, moved, affectionate, surprised, disappointed, sad, angry_frustrated, contempt, anxious_nervous, scared, embarrassed_awkward, confused, neutral

The mixed_more_than_one response option requires the reader to select two or more emotions from the base 14, making response-level multiplicity a first-class response category.

3.4 Response protocol

For each clip, participants (i) watched the curated video window in an embedded player scoped to the specific segment; (ii) read the target-person framing; (iii) selected their primary emotion (or mixed_more_than_one, in which case they selected 2+ constituent emotions); (iv) selected one or more perceptual cues that drove their reading, from a 9-cue taxonomy (facial_expression, tone_of_voice, verbal_content, body_language, situation_context, timing_pacing, others_reaction, something_else, not_sure); and (v) rated their confidence on a 1–5 Likert scale.

Once submitted, a response could not be edited. First impressions are the phenomenon under study.

Session length was capped at 10 clips per session; participants could complete multiple sessions across days.

3.5 Aggregation

For this preliminary descriptive report, analyses operate at the clip level with two core metrics:

- **Modal share:** for each clip, the fraction of readers whose response mentioned the modal (most-frequently-mentioned) emotion, computed as (number of readers mentioning the modal emotion) / (number of readers responding to the clip). Mentions include both single-primary picks and secondary picks made under a mixed_more_than_one primary. Range: 0 to 1.
- **Distinct emotions per clip:** the number of unique emotions that received at least one mention on the clip.

For the H2 descriptive comparison, mean modal share is reported by curator-coded verbal-dependency bucket.

For the multi-emotion-response analysis in Section 4.4, the frequency distribution of number-of-emotions-selected-per-response is reported directly at the response level (not a within-subject comparison in the usual statistical sense — the analytic unit is the individual response, and each reader can contribute multiple responses). The mean confidence rating is computed separately for mixed and single-emotion responses, and the mixed-response rate is reported by confidence rating.

Formal statistical apparatus (Shannon entropy, Krippendorff's α, correlation and t-tests against pre-registered thresholds) is not applied in this preliminary descriptive report and is reserved for a subsequent empirical paper on an expanded corpus.

4. Results

4.1 Sample and stimulus coverage

At the data cutoff, 43 adult participants had produced 442 non-excluded responses on 33 live clips. Per-clip N ranged from 10 to 22, with a median of 11 responses per clip (mean 13.4). All 33 clips are included in the analyses below.

Table 1. Response distribution by wave.

Wave	Clips	Responses	N per clip (range)	Median
Wave 1	11	113	10–11	10
Wave 2	11	132	10–17	10
Wave 3	11	197	17–22	17
All	**33**	**442**	**10–22**	**11**

4.2 Across-reader plurality (H1, descriptive)

On the median clip in this dataset, the most-picked emotion was chosen by 70% of readers (IQR: 50–76%; range: 27%–100%). Roughly five distinct emotions were mentioned on the median clip (IQR: 4–7; range: 1–10). Eight of 33 clips (24%) reached the ≥ 80% modal-share level used here to define a consensus tier; ten of 33 clips (30%) had a modal share at or below 50%, meaning at least half of readers on those clips landed on emotions other than the most-picked one.

Table 2. Descriptive summary of clip-level plurality indicators (N = 33 clips).

Metric	Min	Q1	Median	Q3	Max
Modal share	0.27	0.50	0.70	0.77	1.00
Distinct emotions in use	1	4	5	7	10

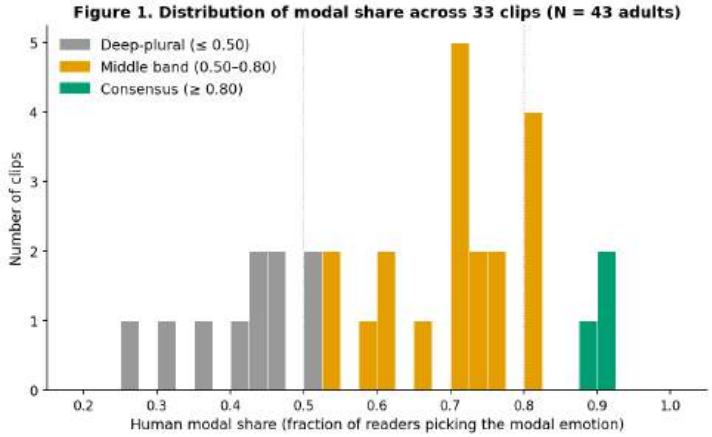


Figure 1. Distribution of modal share across 33 clips, colored by tier.

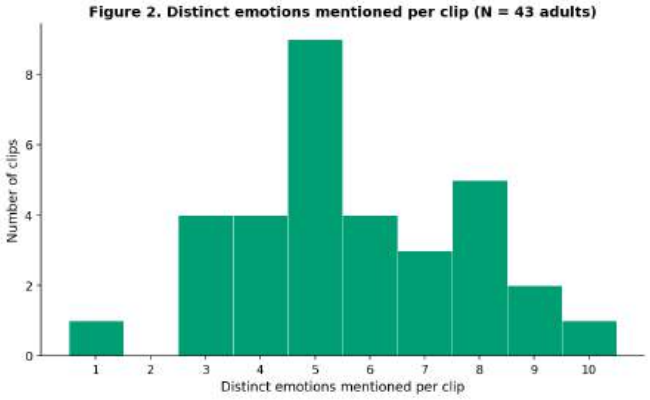


Figure 2. Distinct emotions mentioned per clip across the 33 clips.

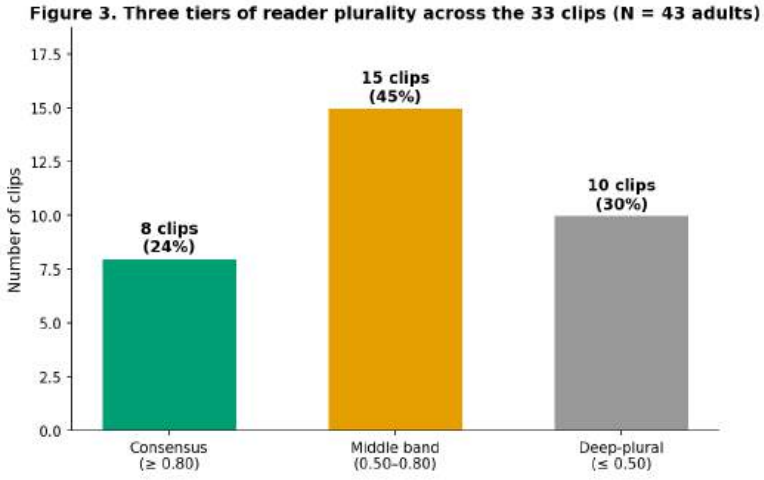


Figure 3. Three tiers of reader plurality across the 33 clips.

Within the observed distribution, three tiers deserve separate attention. On the **eight clips (24%) where modal share reached 0.80 or above** (the *consensus* tier, at the ≥ 80% modal-share threshold used here), human readers behave much as the single-label convention assumes: one interpretation dominates, and a training-corpus curator applying the standard filter would keep those clips. On the **ten clips (30%) where modal share fell at or below the mathematical majority threshold of 0.50** (the *deep-plural* tier), at least half of readers landed on emotions other than the top pick, and the top emotion cannot even be called a majority reading in the strict sense; no single interpretation carries the moment on those clips. Between these two tiers sits **the largest group of clips (15 clips, 45%)**, the *middle band*, where the top emotion is chosen by roughly two thirds of readers but a substantial minority reads the moment differently. These middle-band clips are neither cleanly consensus nor clearly plural, and it is here that the single-label convention is quietest and hardest to notice.

Table 3. Five most-plural clips (lowest modal share).

Clip	N	Modal share	Modal emotion	Distinct emotions
Elemental Movie	11	0.27	Angry/frustrated	6
Kim's Convenience Scene	10	0.30	Proud	8
Love Actually	19	0.37	Disappointed	9
Sarcastic To Each Other	10	0.40	Angry/frustrated	6
The Devil Wears Prada	14	0.43	Confused	6

On the most-plural clip in the collection (Elemental Movie), fewer than a third of readers chose the modal emotion and six different emotions were mentioned. Clips of this shape are the visual counterexamples to the single-label convention: no single reading dominates, and any evaluation that scored an AI as "correct" for picking Angry/frustrated on this clip would be scoring against a modal that only three readers in eleven chose.

4.3 Verbal dependency and agreement: the descriptive comparison (H2)

The pre-registered expectation, motivated by the intuition that dialogue-heavy clips make emotional content more explicit, was that clips with higher verbal dependency would show higher inter-reader agreement, a higher modal share. The descriptive comparison points the other way (Table 4).

Table 4. Mean and median modal share by curator-coded verbal dependency.

Verbal dependency	N clips	Mean modal share	Median modal share	SD
Low	12	0.703	0.707	0.155
Medium	15	0.601	0.588	0.184
High	6	0.608	0.589	0.235

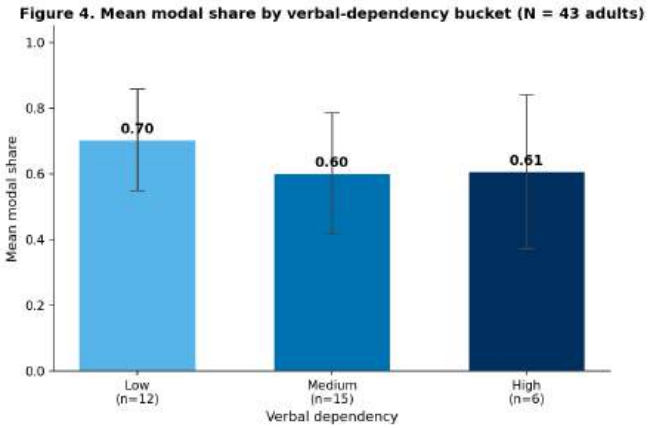


Figure 4. Mean modal share by verbal-dependency bucket, with SD error bars.

Low-verbal clips show the highest mean modal share; medium- and high-verbal clips are essentially tied at the lower end. This is reported here descriptively and without a formal statistical test; the pre-registered correlation test between verbal-dependency ordinal and modal share is reserved for the subsequent empirical paper on the expanded corpus.

The observation, on its own, is informative. Clips heavy on dialogue in this stimulus set tend to depict complex social interactions (sarcasm, awkward apologies, layered emotional exchanges) where the linguistic content is what makes the moment ambiguous rather than what resolves it. Low-verbal clips tend to depict more recognizable emotional moments (laughter, tears, expressions of anger) where the non-verbal signal is shared across readers. The descriptive pattern suggests that verbal *ambiguity*, rather than verbal *amount*, may be the operative predictor of reader disagreement; testing that claim directly would require a new curator-coded variable and is beyond the scope of this paper.

4.4 Multi-emotion responses

A distinct pattern in the data operates at the level of the individual response rather than the level of the clip. **Of 442 responses, 106 (24.0%) explicitly used the mixed response option to select more than one simultaneous emotion.** In other words, 24% of individual responses in this dataset reported reading more than one emotion for the moment the reader had just watched. The mean number of emotions selected per response was 1.42. (The unit of analysis in this subsection is the individual response, not the reader — each reader contributed multiple responses to this pool.)

Table 5. Distribution of emotions selected per response.

Emotions selected in response	Count	% of responses
1 (single emotion)	336	76.0%
2 (dual)	55	12.4%
3 (triple)	31	7.0%

4 (quad)	13	2.9%
5	6	1.4%
6	1	0.2%

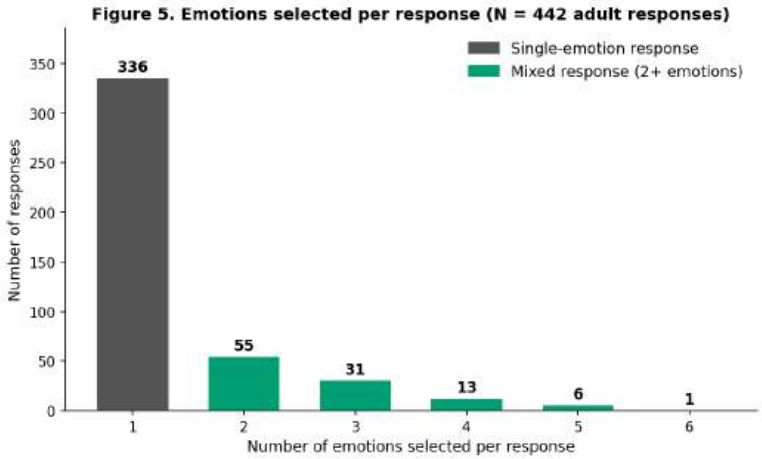


Figure 5. Multi-emotion responses: distribution of emotions selected per response (N = 442 adult responses).

Mixed selection carries nearly the same confidence rating as single-emotion selection in this sample. The mean confidence rating on mixed responses was 4.24 (SD = 0.79, n = 106), while the mean confidence rating on single-emotion responses was 4.44 (SD = 0.73, n = 336), a difference of 0.20 on the 5-point Likert scale. Overall confidence remained high on both response types (mean 4.39 across all responses; 88% of responses at confidence 4 or 5). If mixed responses primarily reflected uncertainty, substantially lower confidence might be expected on mixed responses than on single-emotion responses; the observed 0.20 gap is small and both means sit within the "high confidence" band.

Mixed-response rate by confidence level. A more direct test of whether the plurality pattern is driven by uncertainty compares the rate of mixed responses across confidence ratings. Table 5b reports the breakdown.

Table 5b. Mixed-response rate by confidence rating (N = 442 non-excluded adult responses).

Confidence rating	N responses	Mixed responses	Mixed %
5 (highest)	234	46	19.7%
4	155	41	26.5%
3	46	17	37.0%
2	6	2	33.3%
1 (lowest)	1	0	0.0%

Coarsened splits: at the highest confidence (rating 5), the mixed rate is 19.7% (46 of 234); at lower confidence (ratings 1–4), it is 28.8% (60 of 208). Splitting instead into confidence 4–5 (high) and 1–3 (low) yields 22.4% (87 of 389) versus 35.8% (19 of 53).

Two observations follow. First, the mixed rate does fall as confidence rises, so some of the mixed response on this corpus is associated with reader uncertainty. Second, the effect is bounded: at the highest confidence rating (5), roughly one response in five is still a mixed response. Multi-emotion selection therefore is not fully accounted for by low reader confidence in this sample — a substantial fraction of it persists at maximum confidence, consistent with the interpretation that some readers deliberately report multiple simultaneous emotions when confident that more than one applies.

The response-level pattern complements the clip-level pattern in Section 4.2: both point toward the same observation, that the human response distribution on these 33 clips is not well represented by a single canonical label, but they measure it at different levels of the data.

4.5 Cue attribution patterns

Across all 442 responses (cue field is multi-select, so response-level percentages sum to more than 100%), the perceptual cues participants cited most often were listed in Table 6.

Table 6. Perceptual cues cited across all 442 non-excluded adult responses.

Cue	% of responses citing
Facial expression	70.6%
Body language	52.9%
Tone of voice	52.3%
Situational context	51.4%
Verbal content	44.3%
Others' reaction	13.3%
Timing/pacing	6.6%
Not sure	0.7%
Something else	0.5%

Facial expression is the most-cited single cue, but readers rely on it as the sole cue on a minority of responses; on the typical response, readers cite 2–3 cues, with body language, tone, and situational context each mentioned by roughly half. Something else and Not sure together account for a little over 1% of cue mentions, indicating that participants rarely indicated that their reading fell outside the prespecified cue taxonomy.

5. Discussion

5.1 What the descriptive pattern shows

Across two distinct levels of the data, the observation is the same: the human response distribution on the short social video clips in this corpus is not well represented by a single label, and readers do not converge on a single reading for most clips. Across readers, the median clip in this dataset draws its modal emotion from about seven readers in ten, with roughly five distinct

emotions in play; only one clip in 33 reaches full convergence. At the response level, roughly one in four responses (24%) explicitly reports more than one simultaneous emotion, and does so with confidence ratings nearly identical to those of single-emotion responses. These are not statistical tests; they are direct observations from the response data.

The plurality is not uniform across the 33 clips. On roughly one clip in four, human readers converge tightly enough that the single-label convention approximates their behavior (the consensus tier documented in Section 4.2). On roughly three clips in ten, the top human emotion pick is not even a majority reading (the deep-plural tier). The remaining nearly half fall in the middle band, where the top emotion draws roughly two thirds of readers but a substantial minority disagrees. **Emotion-AI systems evaluated against a single "correct" label on these middle-band clips are being scored against a target that a third of human readers themselves do not share.**

The observations are enough, on their own, to complicate the single-label convention as a description of how humans read this kind of stimulus. If two thirds of readers converge on a modal emotion and the remaining third select emotions that are not random noise (they are, on inspection, coherent alternative readings: a Disappointed clip that a third of readers see as Sad and a fifth as Moved), then a training-data pipeline that assigns a single canonical label per clip is discarding information the human observer distribution contains. Whether that discarded information matters for a given downstream tool is an empirical question, one this paper cannot answer, but one the released dataset is designed to support.

5.2 Verbal content is not the axis originally hypothesized

The descriptive comparison in Section 4.3 pointed in a direction opposite the pre-registered expectation. Low-verbal clips in this corpus showed the highest mean modal share; medium- and high-verbal clips landed noticeably lower and essentially tied with each other. This is reported here as an observation rather than a claim, because the sample sizes at the extremes and the reliance on a single categorical curator code make formal inferential claims premature for this paper. But the observation itself is informative. It suggests that in this stimulus set, the clips that make dialogue the load-bearing channel are also the clips that make the emotional content ambiguous (sarcasm, awkward apologies, layered social exchanges), whereas the clips that make facial and body signals the load-bearing channel are also the clips where those signals are shared across readers. If the pattern generalizes, the operative predictor of reader disagreement is not the *amount* of verbal content but the *ambiguity* of the verbal content. Testing that reframing is a task for future work.

5.3 Two levels of plurality: a summary observation

The core empirical contribution of this paper sits in the convergence between Section 4.2 (across-reader plurality) and Section 4.4 (multi-emotion responses). Two very different measurements, one operating at the clip level, one operating at the response level, point in the same direction, and neither measurement depends on the choice of statistical apparatus. On the

clip level, most clips draw substantive support for multiple emotions from their reader pool. On the response level, roughly a quarter of responses report multiple simultaneous emotions per moment, most at high confidence — the mixed rate falls from 37% at confidence 3 to 20% at confidence 5, but does not disappear, so multi-emotion selection is not fully accounted for by low-confidence hedging. Both patterns are what a constructivist account of emotion perception (Barrett, *How Emotions Are Made*) would predict; both patterns complicate the single-label convention in emotion-annotation practice.

5.4 Implications for emotion-AI training data and evaluation

The dataset released with this paper is response-level, unfiltered, and versioned. Every reading is preserved; no majority-vote reduction is performed at storage time. This is a specific methodological choice: it makes the plurality patterns documented above continuously observable in the released data, rather than smoothed out by a canonical-label step at annotation time.

For downstream emotion-AI training on similar corpora, the descriptive observations suggest two things.

First, the single-label convention adopted by most emotion-recognition datasets does not match the empirical distribution of human readings on short social clips in this sample. Whether that mismatch matters for what a resulting tool can teach a downstream user is an empirical question this paper cannot answer, but the released dataset is designed to support it.

Second, evaluation frameworks that report top-1 accuracy against a single "correct" label are, in the descriptive terms of this paper, evaluating an AI against a target that only about seven readers in ten typically share. A richer evaluation frame — reporting where an AI sits relative to the full reader distribution — recovers information a single-number accuracy metric cannot. The companion paper (Kim, "Where One AI Diverges") takes up that project.

5.5 Limitations

Preliminary-report scope. This paper is deliberately preliminary. Formal hypothesis tests against the pre-registered thresholds (including Shannon entropy over emotion mention distributions, Krippendorff's α as an inter-rater reliability measure, and correlation and t-tests on the H2 and H3 questions) are reserved for a subsequent empirical paper on an expanded corpus (target $N > 100$) that will supersede this preliminary report rather than partition its findings across two publications. The current paper's claims are correspondingly modest.

Sample size. $N = 43$ adult participants is small for population-level claims. Effect sizes should be treated as preliminary; comparisons across sub-groups (verbal dependency, country of origin) are underpowered for firm inferential claims.

Sample composition. Recruitment through personal networks and an international school skews the sample WEIRD-leaning with an internationally-mobile subset. Country-stratified analyses, deferred to future work, will require interpretation under this constraint.

Non-clinical sample. This paper reports on a general adult sample. Claims about autistic or otherwise neurodivergent readers await later project phases with appropriate consent frameworks and partner organizations.

Compact taxonomy. The 14-emotion, 9-cue taxonomy is a deliberate compact set. Finer emotions (awe, nostalgia, relief, and others in Cowen and Keltner's 28) are captured only through the mixed option and free-text responses.

Consent-framework gap. Six minor participants (ages 13–17) registered during the data-collection period but are excluded from the main analyses because verified parental consent was unavailable. See Ethics Statement.

5.6 Future work

- **A subsequent empirical paper on an expanded corpus.** This paper is a preliminary report on the current sample (N = 43 adults). A subsequent empirical paper will extend the sample to N > 100 with additional clips at the low and high verbal-dependency and social-complexity extremes, and will add formal statistical apparatus (Shannon entropy over emotion-mention distributions, Krippendorff's α as inter-rater reliability, correlation and t-tests against the pre-registered H1–H3 thresholds, and the H3 social-complexity analysis, which is not reported here because the extreme buckets contain only six clips each). That paper is intended to supersede, not partition, the present preliminary report.
- **AI comparison.** A companion paper (Kim, "Where One AI Diverges") compares one AI's readings (Gemini 2.5 Pro) to the human distribution documented here, using the same 33-clip corpus. That paper introduces a two-condition methodology comparison to isolate the effect of curator metadata on the AI's reading, and documents a distinct AI-hedging behavior that mirrors the multi-emotion responses observed in the present paper.
- **Cross-cultural analysis.** With per-country N > 20 within-clip, structured cross-cultural comparisons become testable. Current data has SG = 16, KR = 14 participants, still short of per-country per-clip depth.
- **Verbal ambiguity as a distinct variable.** The observation in Section 4.3 (that verbal *amount* points the wrong way, but verbal *ambiguity* may be the real predictor) invites a new curator-coded variable and a fresh data collection to test it directly.
- **Human information architecture as its own variable.** This paper's human protocol held one information bundle constant: video plus a target-person description. A parametric extension could vary the cultural, relational, and search-augmented context available to a reader at reading time, testing whether inter-reader plurality is itself a function of information architecture.
- **Phase 2.** Validated teaching materials grounded in the plural-reading dataset are the eventual project scope.

6. Data and Code Availability

- **Dataset:** Response-level dataset (adult participants) available at https://www.mindlenslab.org/live-findings. Minor-participant responses are not analyzed or released in this report (see Ethics Statement).
- **Pre-registration:** v0.1 pre-analysis protocol at https://www.mindlenslab.org/research/hypotheses.
- **Analysis code:** Python analysis pipeline for the descriptive statistics reported here is released at https://github.com/evelyn-mindlens/preprint-analysis under the MIT license.
- **License:** Data released under CC BY 4.0. Code released under MIT.

Recommended citation: Kim, Evelyn. "Plural Emotion Readings in Short Social Video Clips: A Descriptive Study." *MindLens Lab Preprint*, 2026, https://mindlenslab.org.

7. Ethics Statement

Consent framework. All participants self-consented on a versioned informed-consent document (v1, deployed with the study platform in May 2026) presented at registration. The consent instrument required each participant to affirm the following: (i) that they were 13 years of age or older; (ii) that they understood the study consisted of watching short publicly-available video clips (10–30 seconds) and selecting emotion labels; (iii) that their pseudonymous responses would be stored, analyzed, and released as an open academic dataset under a CC BY 4.0 license; (iv) that no identifying information (name, address, contact details beyond an access email) would be retained in the analyzed dataset; (v) that no compensation was offered; and (vi) that they retained the right to withdraw and request data deletion at any time via contact@mindlenslab.org. The full v1 consent text is included as Supplementary Material S4.

Minor participants. During this data-collection period, 6 of the 49 non-test-account participants who submitted at least one response were minors (ages 13–17). Because verified parental consent was unavailable for these six participants, their responses were excluded before analysis and are not analyzed or released in this report. The study platform's parental-consent verification workflow is implemented at the schema and application level and will be gated at session entry for future data-collection windows.

Data handling. All data is pseudonymous. Participants' email addresses are used only for account access and result delivery and never appear in analysis files. Data deletion is available on request via contact@mindlenslab.org. No compensation was offered. No deception was used.

IRB status and applicability. This study was conducted as an independent citizen-science project by an unaffiliated high-school researcher and was not conducted under the auspices of any US-federally-funded research institution; it therefore falls outside the jurisdiction of any Institutional Review Board. Under the U.S. Common Rule (45 CFR 46), the study design would additionally qualify as exempt research under §46.104(d)(2)(i) — research involving benign behavioral interventions with adult participants in which the information obtained is recorded such that the identity of subjects cannot readily be ascertained — on the following grounds: (a) the analyzed and released sample is restricted to adults aged 18 or older; (b) the task

is non-invasive, consisting of watching short publicly-available video clips and selecting emotion labels; (c) participants provided documented informed consent (v1, above; Supplementary S4); (d) no personally identifying information is retained in the analyzed dataset; (e) participants received no compensation, removing coercion concerns; and (f) participants may withdraw and request data deletion at any time. These considerations are documented openly on the recruitment page and in the consent framework, and are surfaced here to make the study's ethics posture explicit rather than implicit.

8. Author Contributions

- **Evelyn Kim:** Conceptualization, Methodology, Data curation, Investigation, Formal analysis (descriptive), Writing – original draft, Writing – review and editing, Visualization, Project administration.

9. Competing Interests

The author declares no competing interests. MindLens Lab is a non-commercial research project. No external funding was received for this work.

10. Acknowledgments

The author is grateful to every participant who contributed a reading; their responses are the substance of this work. Thanks are also due to the Singapore American School community for recruitment support and the academic environment that shaped the originating paper, and to Zhi Hao Lim (PhD candidate in Economics, Cognition and Decision Lab, Columbia University) for methodological guidance on study design and preprint drafting. Sean Kim (parent) provided site planning and AI-assisted development support of the MindLens Lab platform.

AI tools (Claude Sonnet, via Anthropic's Cowork mode) were used to support code implementation, data checking, descriptive-statistics computation, and manuscript editing. The author designed the study, curated the stimuli, collected the data, reviewed all analytical outputs, interpreted the findings, and took responsibility for the final manuscript. This work rests on the author's originating paper (Kim, "Artificial Intelligence") and its mission.

Works Cited

Abbas, Zayn Jameel, and Stacey D. Scott. "The Pluralistic Nature of Emotion: Human and Machine Interpretations of Textual Emotional Content." Proceedings of the 31st International Conference on Intelligent User Interfaces (IUI '26), ACM, 2026, https://doi.org/10.1145/3742413.3789161.

Baron-Cohen, Simon. "Autism: The Empathizing–Systemizing (E–S) Theory." Annals of the New York Academy of Sciences, vol. 1156, no. 1, 2009, pp. 68–80, https://doi.org/10.1111/j.1749-6632.2009.04467.x.

Barrett, Lisa Feldman. How Emotions Are Made: The Secret Life of the Brain. Houghton Mifflin Harcourt, 2017.

Barrett, Lisa Feldman, et al. "Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements." Psychological Science in the Public Interest, vol. 20, no. 1, 2019, pp. 1–68, https://doi.org/10.1177/1529100619832930.

Cowen, Alan S., et al. "Mapping 24 Emotions Conveyed by Brief Human Vocalization." American Psychologist, vol. 74, no. 6, 2019, pp. 698–712, https://doi.org/10.1037/amp0000399.

Cowen, Alan S., and Dacher Keltner. "What the Face Displays: Mapping 28 Emotions Conveyed by Naturalistic Expression." American Psychologist, vol. 75, no. 3, 2020, pp. 349–364, https://doi.org/10.1037/amp0000488.

Ekman, Paul. "An Argument for Basic Emotions." Cognition and Emotion, vol. 6, nos. 3–4, 1992, pp. 169–200, https://doi.org/10.1080/02699939208411068.

Golan, Ofer, et al. "The 'Reading the Mind in the Voice' Test–Revised: A Study of Complex Emotion Recognition in Adults with and without Autism Spectrum Conditions." Journal of Autism and Developmental Disorders, vol. 37, no. 6, 2007, pp. 1096–1106, https://doi.org/10.1007/s10803-006-0252-5.

Harms, Madeline B., et al. "Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies." Neuropsychology Review, vol. 20, no. 3, 2010, pp. 290–322, https://doi.org/10.1007/s11065-010-9138-6.

Kim, Evelyn. "Artificial Intelligence in Emotional Intelligence Training for Autism." Curieux Academic Journal, July 2025 (Part 2 Issue), pp. 394–410.

Kim, Evelyn. "Where One AI Diverges from Plural Human Readings on Short Social Video Clips." MindLens Lab Preprint, 2026, https://mindlenslab.org. Forthcoming on Zenodo.

Landowska, Agnieszka, et al. "Automatic Emotion Recognition in Children with Autism: A Systematic Literature Review." Sensors, vol. 22, no. 4, 2022, article 1649, https://doi.org/10.3390/s22041649.

Latif, Siddique, et al. "AI-Based Emotion Recognition: Promise, Peril, and Prescriptions for Prosocial Path." arXiv, 2022, arXiv:2211.07290, https://doi.org/10.48550/arXiv.2211.07290.

Mollahosseini, Ali, Behzad Hasani, and Mohammad H. Mahoor. "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild." IEEE Transactions on

Affective Computing, vol. 10, no. 1, 2019, pp. 18–31, https://doi.org/10.1109/TAFFC.2017.2740923.

Munafò, Marcus R., et al. “A Manifesto for Reproducible Science.” Nature Human Behaviour, vol. 1, no. 1, 2017, article 0021, https://doi.org/10.1038/s41562-016-0021.

Nosek, Brian A., et al. “The Preregistration Revolution.” Proceedings of the National Academy of Sciences, vol. 115, no. 11, 2018, pp. 2600–2606, https://doi.org/10.1073/pnas.1708274114.

Russell, James A. “A Circumplex Model of Effect.” Journal of Personality and Social Psychology, vol. 39, no. 6, 1980, pp. 1161–1178, https://doi.org/10.1037/h0077714.

Silvera-Tawil, David, et al. “Socially-Assistive Robots to Support Learning in Students on the Autism Spectrum: Investigating Educator Perspectives and a Pilot Trial of a Mobile Platform to Remove Barriers to Implementation.” Sensors, vol. 22, no. 16, 2022, article 6125, https://doi.org/10.3390/s22166125.

Uljarevic, Mirko, and Antonia Hamilton. “Recognition of Emotions in Autism: A Formal Meta-Analysis.” Journal of Autism and Developmental Disorders, vol. 43, no. 7, 2013, pp. 1517–1526, https://doi.org/10.1007/s10803-012-1695-5.

Wang, Yan, et al. “A Systematic Review on Affective Computing: Emotion Models, Databases, and Recent Advances.” arXiv, 2022, arXiv:2203.06935, https://doi.org/10.48550/arXiv.2203.06935.

Zhang, Qing, et al. “Facial Emotion Training as an Intervention in Autism Spectrum Disorder: A Meta-Analysis of Randomized Controlled Trials.” Autism Research, vol. 14, no. 10, 2021, pp. 2169–2182, https://doi.org/10.1002/aur.2565.

Zijderveld, Gabi. “The World’s Largest Emotion Database: 5.3 Million Faces and Counting.” Affectiva Blog, 24 Apr. 2018, https://blog.affectiva.com/the-worlds-largest-emotion-database-5.3-million-faces-and-counting.

Shell Shock Post World War I: Challenging and Encountering Masculinity
By Talia Fine

Introduction

Before World War I, the United States Military had little to no knowledge or experience treating soldiers with mental issues, yet the field rapidly gained relevance as a result of the war. Although WWI introduced major technological advancements in firepower, tanks, and aircraft, one of the largest changes was the increased use of large-scale explosions (Vergun). Combined with the war's expanded global scale, these widespread explosions began to affect soldiers mentally, instead of only physically--a new phenomenon for military and general society.

Over the course of WWI, approximately 15,000 to 76,000 American soldiers suffered from a mental illness coined by British psychologist Charles Samuel Myers in 1915 as shell shock (Stagner, The National Archives). Although the meaning of the term has developed over time, the current definition is, "A disorder identified in soldiers…attributed to exposure to shell-fire and characterized by severe anxiety and other psychological disturbances" (Oxford English Dictionary). Historian Paul R. Deslandes defines masculinity as built "on notions of proficiency, competence, intellectual and psychological as well as physical fortitude, and unimpeachable character" (Paul R. Deslandes 124-125) Before WWI, masculinity had been based on economic stability, but at the turn of the century, historian Eric J. Segal emphasized the role of a "disciplined and fortified male body." (Eric J. Segal) These definitions, which were both influenced by the Anglo-American movement of Muscular Christianity that was growing throughout the US, exhibit how important mental strength and resilience were to masculinity at the time (Dowland 356). The emergence of shell shock exposed what contemporaries perceived as the fragility of traditional masculinity and led the soldiers, military leaders, and physicians to develop new ways of addressing mental health challenges in efforts to preserve ideals of strength, resilience, and military effectiveness.

Notably, this paper does not argue that all soldiers understood shell shock primarily through masculinity; rather, it examines how militaries, medical professionals, media, and veterans themselves increasingly interpreted shell shock through prevailing ideas of masculine strength and resilience.

From Hysteria to Shell Shock: The Gendering of Psychological Trauma

Before World War I, diagnoses of hysteria were strongly gendered and were disproportionately associated with women. French neurologist Jean-Martin Charcot believed that hysteria was caused by an "unknown internal injury affecting the nervous system." (McVean). Sigmund Freud, one of Charcot's medical students, was an Austrian neurologist who later founded psychoanalysis and conceived many theories on psychology. Freud's belief was that a woman suffered from hysteria because of her repression of male sexuality (McVean). In centering the idea of hysteria around the absence of male features, Freud immediately created a harsh stigma of femininity and weakness around the term. No man would want to be called

hysterical, as the term was associated with the female body and a lack of strength. For example, Meyers did not want to "stigmatize the British soldiers as hysterical," explaining how associating a soldier with mental illness subsequently tied them to a lack of masculinity because of the stigma that Freud had created.

France and Britain entered WWI in 1914, two years before the United States officially joined the war effort. During these years, the two countries and the US, which quickly realized the relevance of the research and the potential impact on soldiers, made significant developments in their knowledge of shell shock. Britain originally believed that shell shock was a type of physical injury that created internal injuries. In 1915, an article in the *Los Angeles Times* written by a British physician described shell shock as a 'psychic knockout' caused by the air pressure from the explosions (Stagner 258). By 1916, however, as Britain and the US continued to investigate shell shock, they realized that the illness was actually a type of psychological trauma caused by seeing the horrors of war. In an article published in *The Living Age* on September 2, 1916, a British soldier stated that the war caused "trench nerves" (Stagner 258). W.R. Houston, an American psychologist who explored shell shock in European soldiers, wrote an article on March 25, 1917, in the *New York Times*. In his piece, he stated that the war created an emotional strain on both mentally stable and unstable soldiers: "This wreckage of men's souls…seemed to me to mirror more vividly the horror of war" (Houston). Houston's realization marked a major shift in recognizing that shell shock represented a psychological injury rather than a physical one. The updated understanding led doctors to classify soldiers under the same general mental illness category previously reserved for women diagnosed with hysteria. The expansion of the "hysteria" diagnosis to include male war neurosis contributed to the transformation of a soldier's trauma into a debated topic of masculinity; as a result, the strength and fortitude of soldiers diagnosed with shell shock immediately began to be challenged.

Heroism and Masculinity: Reframing Shell Shock in Wartime Culture

Although the association with hysteria threatened to diminish soldiers' masculinity, soldiers with shell shock were initially viewed as heroes because they continued to fight despite suffering from shell shock. Through public media, shell shock thus became separated from the stigma of hysteria. For example, the *Stars and Stripes* AEF newspaper, published in France, validated shell shock as an injury sustained from war, and distinguished the illness from civilian mental disorders such as hysteria (Stars and Stripes 2). Additionally, an article in *Scribner* wrote that war was the only means that could cause shell shock, and that a soldier who suffered from the illness was not weak, but rather a hero who fought despite the dangers of war (Stagner 261). Because of their ability and willingness to sacrifice mental capacity for the war, soldiers who were diagnosed with shell shock were thought of as those who had fought most valiantly in battle. As a result, the masculinity of soldiers with shell shock was upheld despite the initial linkage to hysteria and femininity, and many soldiers received special praise for continuing to fight despite their condition (Stagner 260). So, even though the stigma of mental illness, which is associated with weakness, challenged the masculinity of a soldier, public media originally

reconstructed the idea and provided societal support for the soldiers. As the heroic connotations were directly related to the war, continuing to fight became a way that soldiers could uphold their masculinity and receive praise for their strength--in other words, to reaffirm their masculinity even though they suffered from mental health challenges.

Proving Manhood: Military Recruitment and Psychological Screening

For decades, the act of enlisting in the US military was viewed as a heroic act of someone who was willing to sacrifice themselves for the betterment of the US. The US used creative posters to recruit new soldiers at the beginning of WWI. One of these posters, made in 1917, titled “Let This Grand Old Bell Proclaim Another Victory," was produced with text: “Enlist Now!/Your Forefathers did in '76/Your Grandfathers did in '61/Your Fathers did in '98/Now You Do It in '17” (NYSL). By directly tying one's duty to one's family’s history, the poster created intense societal pressure, suggesting that to be a worthy descendant or figure in society, one had to fight. Laura Brey created another 1917 recruiting poster titled “Enlist: On Which Side of the Window Are You?” The image depicts a well-dressed man standing in a dark room watching troops with an American flag march outside, representing the choice between staying in the shadows of safety as a coward or joining the bright world of soldiers fighting to uphold American values despite the personal dangers to themselves. The poster was meant to shame men who were able but unwilling to enlist in the war (Brey). As a result of these posters, a societal expectation was created in America; enlisting was essential to proving a man's strength. Following media publication of the trauma and dangers of shell shock from the experiences of soldiers in other countries, enlisting was increasingly perceived as a brave act. Therefore, joining the war became a chosen way that a man could independently assert and uplift his own masculinity.

By the time the United States joined the war in April of 1917, researchers, psychologists, and governmental organizations already had substantial knowledge of shell shock. As a result, the US implemented pre-screening programs for a soldier's mental capacity to determine the likelihood of a soldier suffering from shell shock. One of these programs was called Army Alpha, which evaluated US military recruits for intellectual and mental ability (AMEDD). These programs tried to preemptively remove soldiers with mental instabilities who would be at a higher risk of shell shock. Because serving in the military was believed to be the responsibility of a man, their mental state became seen as a sign of strength, and soldiers who were not approved to be in combat situations were thought of as not fulfilling their responsibility as men in US society. In doing so, the US challenged the masculinity of many soldiers who wanted to display their strength and exhibit their place in society as men.

Restoring the Soldier: Treatment, Recovery, and the Burden of Masculinity

The treatment for shell shock at the time also undermined the core of a soldier's masculinity. The US military followed Forward Psychiatry (PIE), which prioritized proximity, immediacy, and expectancy, as the goal of the US military was to have the soldier return to the

front lines as quickly as possible. Initial care for shell-shocked soldiers was provided by the Regimental Medical Officer, who was positioned in close proximity to the front lines (The Western Front Association). Mild cases were treated in local specialized hospitals, and the main treatment was psychodynamic intervention, which used tools like electric shock to cure the physical impact of shell shock, such as tremors (Alexander). It can be inferred that, because doctors thought of shell shock as a mental health issue and therefore a form of non-masculine cowardice that challenged masculinity, they used these harsh, physically brutal forms of treatment to return a soldier to his previous strength and masculine personality. Thus, the military and society thought of those who were unable to recover as weak and feminine, relating back to the original connection to hysteria.

Soldiers with severe cases of shell shock were taken out of battle and moved to neurological centers and hospitals farther from the front lines. U.S. Army Base Hospital No. 117, located in La Fauche, France, was one of the premier hospitals during WWI for treating shell shock. Although the hospital was focused on returning soldiers to their previous mental state, the main goal of the hospital was to restore soldiers to their full military ability so that they could return to fighting (AMEDD). In other words, the societal expectation, ingrained into the soldiers, was that a true soldier would continue to fight even after experiencing shell shock. Returning home rather than returning to the front could be viewed as a weakness and a failure of masculine duty and mental fortitude, because that was only for those men who could not maintain the mental strength needed for a real man who could fight and fulfill their societal responsibility.

At the time, if a soldier was unable to recover to go back to the front lines, he would be sent to a specialist neurological center or base hospital. From there, only the most severe cases of shell shock would be approved to return home if they required long-term care or were deemed a chronic mental illness (Alexander). Because soldiers were expected to return to war, returning home was sometimes seen as a personal weakness rather than the result of an injury. As a result, many soldiers were pushed back to the front--and pushed themselves back to the front--to avoid being perceived as weak despite needing additional treatment, which contributed to higher rates of permanent disability among soldiers.

Once a soldier returned home, the impact on their masculinity continued. At home, soldiers did not speak of the shell shock they experienced, partly from trauma, and partly because telling others that they needed to come home and not fulfill their societal duty because of a mental illness undermined their masculinity. Reese Melvin Russell, an American soldier who suffered from shell shock, kept a diary during and after the war. His daughter Frances wrote that after the war, “my father could no longer relate to the world he left behind” (Russell). According to Frances, he never spoke to anyone about what he saw during the war, and whenever the war was brought up in history lessons, he would avoid discussion in an attempt to conceal the emotional impact the war had caused for him. Russell’s silence suggests the social pressures many veterans faced when discussing psychological trauma, pressures closely tied to contemporary expectations of masculinity and fortitude.

Conclusion: Shell Shock and the Reconstruction of Masculinity

Less widely known than many of the other aspects of WWI, the new phenomenon of shell shock hit at the very core of the masculinity of the brave soldiers fighting for their country, causing them to confront these challenges and to come up with ways to retain some of their masculinity. Initially viewed as heroes for fighting despite suffering the clear impact of shell shock, those soldiers who were not able to resume fighting became seen as weak and feminine, unfortunately forcing them to continue to confront the stigma of mental illness throughout their lives in silent, often subtle ways.

Works Cited

Alexander, Caroline. “The Shock of War.” Smithsonian Magazine, Sept. 2010, https://www.smithsonianmag.com/history/the-shock-of-war-55376701/. Accessed 2 Apr. 2026.

“A Case of Shell Shock.” Stars and Stripes, 14 June 1918, p. 2.

AMEDD Center of History & Heritage. “Hospital for War Neuroses (Base Hospital No. 117).” https://achh.army.mil/history/book-wwi-neuropsychiatry-section2chapter4. Accessed 25 Mar. 2026.

Brey, Laura. “Enlist: On Which Side of the Window Are You?” Library of Congress, 1 Jan. 1970, https://www.loc.gov/pictures/item/98503178/. Accessed 2 Apr. 2026.

Deslandes, Paul R. Oxbridge Men: British Masculinity and the Undergraduate Experience, 1850–1920. Indiana University Press, 2005.

Dowland, Seth. “War, Sports, and the Construction of Masculinity in American Christianity.” Religion Compass, vol. 5, no. 7, July 2011, pp. 354–67. https://doi.org/10.1111/j.1749-8171.2011.00278.x. Accessed 2 Apr. 2026.

Houston, W. R. “War’s Amazing Effect on Nerves of Soldiers; Distinguished Neurologist Tells from Personal Observation of Remarkable Cases of Hysteria That Have Resulted at the Front.” The New York Times, 25 Mar. 1917, https://www.nytimes.com/1917/03/25/archives/wars-amazing-effect-on-nerves-of-soldiers-distinguished-neurologist.html. Accessed 2 Apr. 2026.

Houston, W. R. “War’s Amazing Effect on Nerves of Soldiers; Distinguished Neurologist Tells from Personal Observation of Remarkable Cases of Hysteria That Have Resulted at the Front.” The New York Times, 25 Mar. 1917, https://www.nytimes.com/1917/03/25/archives/wars-amazing-effect-on-nerves-of-soldiers-distinguished-neurologist.html. Accessed 2 Apr. 2026.

“Hun Attack Smashed by Yankee Defenders.” Stars and Stripes, 26 Apr. 1918, p. 1.

McVean, Ada. “The History of Hysteria.” Office for Science and Society, 31 July 2017, https://www.mcgill.ca/oss/article/history-quackery/history-hysteria. Accessed 2 Apr. 2026.

Oxford English Dictionary. “Shell Shock, N. Meanings, Etymology and More.” Oxford English Dictionary, https://www.oed.com/dictionary/shell-shock_n. Accessed 25 Mar. 2026.

Segal, Eric J. “Norman Rockwell and the Fashioning of American Masculinity.” The Art Bulletin, vol. 78, no. 4, Dec. 1996, pp. 633–46. https://www.jstor.org/stable/3046212.

Stagner, Annessa C. “Healing the Soldier, Restoring the Nation: Representations of Shell Shock in the USA during and after the First World War.” Journal of Contemporary History, vol. 49, Apr. 2014, pp. 256–74. https://www.jstor.org/stable/43697299.

Russell, Reece Melvin. “Reece Melvin Russell Collection.” Library of Congress, 1917. https://www.loc.gov/item/afc2001001.18879/

The National Archives. "'Shell Shock' Cases." The National Archives, 9 Sept. 2020, https://www.nationalarchives.gov.uk/education/resources/medicine-on-the-western-front-part-two/shell-shock-cases/. Accessed 25 Mar. 2026.

The Western Front Association. "Shell Shock: Genesis and Outcomes in the Great War." The Western Front Association, 22 May 2008, https://www.westernfrontassociation.com/world-war-i-articles/2008/may/shell-shock-genesis-and-outcomes-in-the-great-war/. Accessed 25 Mar. 2026.

Vergun, David. "Americans Answer Call to Arms during World War I." U.S. Department of War, 9 Nov. 2021, https://www.war.gov/News/News-Stories/Article/Article/2839338/americans-answer-call-to-arms-during-world-war-i/. Accessed 25 Mar. 2026.

"World War I Posters (SC23360) – United States (Recruiting Posters)." New York State Library, https://www.nysl.nysed.gov/msscfa/sc23360-us-recruit. Accessed 25 Mar. 2026.

How Do Historical Medical Injustices Affect Trust in Gene Therapy Among Minority Communities? By Kayla Nicole Ofosu

Abstract

Gene therapy is a cutting-edge medical innovation that involves altering the genes inside a person's cells to treat or prevent disease. With its innovative techniques, this process has the potential to revolutionize healthcare and the field of medicine. Gene therapy has been used to lessen the symptoms of and treat immunodeficiency disorders, has been tested in various cancer trials, and has been viewed as a potential treatment for HIV and AIDS (Thompson 1310). Yet, gene therapy is not widely available to the public; however, doctors and scientists evaluate cases to determine if gene therapy would be a suitable treatment option. The trust of the public, especially historically marginalized groups, is important to its success. Minority communities are populations that have encountered systemic limitations to healthcare access and have historically been subjected to unethical medical practices. Health disparities are unfair and avoidable differences from political, social, and economic standpoints that result in varying health outcomes between populations. These communities can include African Americans (affected by the Tuskegee Syphilis Study and the case of Henrietta Lacks), Indigenous communities (impacted by forced sterilizations), and Hispanic/Latino groups (who have experienced disparities in medical research, inclusion, and healthcare equity) (Brandt 21–29; Carpio 40–53; Ndugga et al.).

Historical Context of Medical Injustices

Ethical violations, lack of consent, and harm to marginalized communities often mark medical experimentation on minorities. The Tuskegee Syphilis Study, for example, took place from 1932 to 1972. During this experiment, Black men who had contracted syphilis were left untreated to study disease progression. The researchers deceived the men into believing they were receiving care and expert treatment for "bad blood" but never disclosed their diagnosis or intended to cure them (Brandt 21–29). Furthermore, Henrietta Lacks was a Black woman whose cancer cells were taken without her knowledge in 1951 at Johns Hopkins Hospital. Her cells, later named HeLa cells, were the first human cells to continuously grow and reproduce outside the body. This made them valuable for biomedical research (Office of Science Policy). These cells contributed to groundbreaking medical advancements, including the polio vaccine, cancer treatments, AIDS research, and even COVID-19 studies (Office of Science Policy). However, for decades, Lacks and her family were not informed or rewarded for the medical research Henrietta Lacks created. This highlights the ethical issues in medical research and exploitation of minorities in medical experimentation, which have caused mass distrust in medicine.

Minority Communities' Perceptions of Modern Medicine

Everyone holds differing views within medicine; from marginalized communities to diverse backgrounds, the lack of reliance or trust creates an inevitable barrier to proper

healthcare. Staff Writer—an author of a health equity article published by Harvard T.H. Chan School of Public Health—records that women, members of the LGBTQ+ community, and disabled persons are more likely to distrust their healthcare providers (Harvard T.H. Chan School of Public Health). Discrimination, economic status, and historical factors are said to contribute to how modern medicine is understood. A statement from Bazargan, Cobb, and Assari's research, published by the National Institutes of Health, declares, "An extensive review of the literature revealed that among minority groups, patient mistrust is linked to a lack of commitment to treatment plans, negative health behaviors, and undesired health outcomes" (Bazargan et al. 5). During COVID-19, health care policies demonstrated structural racism within the United States government. African Americans, Latinos, Alaska Natives, and American Indians were among the groups with the highest hospitalization and mortality rates between 2020 and 2023. Authors Ruqaiijah Yearby, Brietta Clark, and José F. Figueroa explain in another article, "Racism includes a complex array of social structures, interpersonal interactions, and beliefs by which the group in power categorizes people into socially constructed 'races' and creates a racial hierarchy in which racial and ethnic minority groups are disempowered, devalued, and denied equal access to resources" (Yearby et al. 187). For example, the federal government enacted the Hill-Burton Act (Hospital Survey and Construction Act) in 1946, which permitted equally available health care facilities yet allowed states to build separate facilities (Yearby et al. 187–94).

Furthermore, insurance coverage is a steadily rising issue amongst minority communities. "As of 2019," concerning health insurance, "58 percent of Americans were covered by employer-sponsored health insurance, with 66 percent of White workers covered by this insurance compared with 47 percent of Black, 43 percent of Latino, and 37 percent of American Indian and Alaska Native workers" (Ahmadi and Murdock). Even though policies, such as the Affordable Care Act, for example, were meant to target inequalities and to offer benefits, low-wage ethnic minority workers continue to suffer. Those who are poorly insured or are not eligible for Medicaid meet a "roadblock" and often fall into a coverage gap, as government administrators debated whether to expand Medicaid (Ahmadi and Murdock).

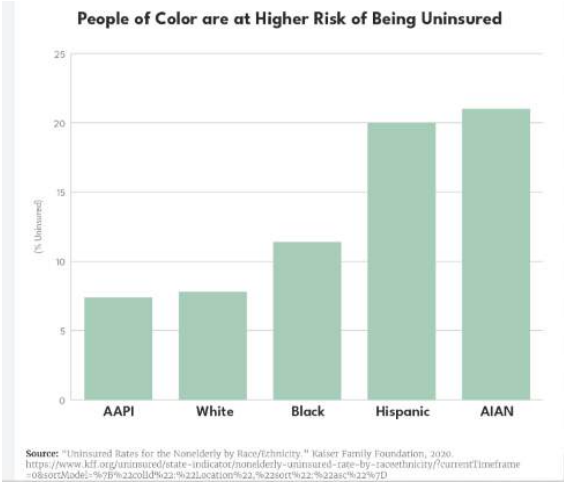


Fig. 1. (Ahmadi and Murdock)

These statistics demonstrate inequalities in the healthcare system and help to explain why some minority communities may be hesitant to participate in clinical trials. Culture, which consists of languages, styles, beliefs, and attitudes, often hinders exposure to new treatments and studies among minorities. Religion and background often overlap with an individual's understanding of the overview, duration, severity, and prevention of an illness. Based on recent cross-sectional studies and surveys, patients who demonstrated a firm foundation in Traditional Chinese Medicine were less likely to adhere to medication(s) for chronic conditions. With a spiritual background, some patients from Ghana even risked forfeiting their medication while anticipating divine intervention (Shahin et al. 1019–35).

Usually causing hesitancy, previous experiences with trauma, discrimination, and narrow representation can impact the effort to participate in such trials. In LGBTQ+ communities, individuals face fears of stigma, restricted health benefits, and institutional bias in healthcare (Lampe et al.). LGBTQ+ persons, specifically Transgender Women, African Americans, and Hispanic/Latino men, usually have a greater risk for HPV/HIV infections or genital cancers. Yet, based on a survey, "33% of respondents who had seen a primary care provider reported negative experiences in the healthcare environment related to their transgender identity, and almost one quarter of all respondents said they refrained from seeking medical care at least once in the preceding year due to fear of being mistreated for their gender identity" (*Second State of LGBTQ Health™ National Survey*). Complications with physical health, access to proper care, clinical research, and even behavioral health contribute to this gap (CDC; National Coalition for LGBTQ+ Health).

Health behaviors, which are described as the ability for individuals to actively maintain good health, can be affected by a wide array of factors including age and gender; however, modifiable factors—such as culture and religion—clearly create a lasting effect on healthcare but can be appropriately amended. These beliefs carry an inescapable burden that hinders patients, especially within minority communities, from receiving appropriate treatment. Scientists and physicians, though, offer numerous techniques, including improving communication, providing support resources, administering sensitive care, and even assessing patients thoroughly through psycho-educational opportunities, that may "break the separation" (Shahin et al. 1019–35).

Gene Therapy and Minority Communities

Despite these prevalent barriers, there are recorded benefits of utilizing gene therapy techniques to assist minority communities. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is known as a Nobel Prize-winning technology to significantly impact the prevention of human disease by gene editing. However, minority communities—who are often prone to endemic illness or genetic diseases—also suffer from a lack of access to such resources and their advantages (Subica 874–82). CRISPR, though a state-of-the-art invention, carries concerns circulating around health equity, as minority populations are usually excluded from clinical trials. According to research analyzed by Andrew M. Subica, "[...] as most participants in genomic studies are of European ancestry, suggesting a future patterning in which minority

populations will fail to benefit equally from CRISPR advances for improving health" (874). By modifying, inserting, and cutting DNA with molecular scissor proteins (Cas9), Scientists can effectively manipulate DNA to further examine the development of human diseases (Subica 874–82). Is the process just, fair, and equitable? Sickle cell anemia, which affects persons of African, Middle Eastern, South Asian, and Central/South American descent, is caused by a single-point mutation in the B-globin gene. This gene is located on chromosome 11 and codes for beta-globin, a key component of hemoglobin. Technology, such as CRISPR, has allowed scientists to accurately detect these point mutations, while regulating genetic factors that worsen certain disadvantages in an environment (Abraham et al.).

Furthermore, CRISPR strengthens the foundation for virus surveillance, food insecurity, and vector-borne disease control. Minority communities, which disproportionately suffer from viral diseases (HIV, Mpox, SARS-CoV-2) that are difficult to track, can benefit from CRISPR-based surveillance programs. These assays can diagnose various viruses rapidly and efficiently in the public to mitigate further community outbreaks. Livestock and crops, which are genetically modified by CRISPR to yield a beneficial trait, may also prove to be an advantage for minority populations worldwide. Additionally, CRISPR can be used to address potential drivers of diseases. Vectors are responsible for nearly 17% of all infectious diseases, including 700,000 deaths worldwide. By "deleting" certain genes in a population of such vectors, CRISPR can block mosquitoes, for example, from transmitting malaria parasites (Subica 874–82).

However, the noted benefits of gene therapy do not eliminate the economic, ethical, and cultural boundaries surrounding its use. In a previous infamous case, the blood samples donated by Havasupai Tribe members for a study around the risk for diabetes were used by Arizona State University researchers for studies on "tribal migration, mental disorders, alcoholism, and interbreeding." As the studies severely violated cultural myths and beliefs, tribal members sued the University and won a financial settlement for the misuse of their genetic data. Not only does this reveal the need for transparent communication within the medical field, but also the necessity for minorities to actively discuss their opinions in community-based research. Concerning the financial statement, gene therapies range from $2.1 million to $3.5 million, with speculations that SCA (Sickle Cell Anemia) therapy could exceed these costs. Insurance coverage, reimbursement issues, and copay burdens further limit the scope for minorities, as these therapies are out of reach for the economically and socially disadvantaged in society. There then becomes a need to increase community awareness and investment in resources, as low availability and funding for efficacious therapies become a greater concern (Subica 874–82; Ndugga et al.).

Addressing the Trust Gap

One of the biggest and most important steps towards rebuilding trust in medical research among minority communities is amplifying the need for diversity in clinical trials. Numerous efforts have been made to diversify clinical trial populations by including more African-Americans, Asian-Americans, and Latinos (Beasley 16). Including these populations in

clinical research and clinical trials is essential to eliminating the discrepancies that can occur when research does not encompass a diverse population. Another key factor in rebuilding trust in minority communities is transparency in research practices. Studies have shown that there is a relationship between transparency and accountability (Zúñiga et al. 1). When research practices and methodologies are more transparent, doctors, scientists, and researchers are more likely to take accountability for their actions, which would increase the trust of minorities in science. In countries like Nigeria, where sickle cell anemia is more prevalent and would greatly benefit from gene therapy, equitable access to medicine is unheard of. Expanding access to medicine is crucial to eliminating and reducing the effects of diseases like sickle cell (Abraham et al.; Ndugga et al.).

Conclusion

After all this, the following can be concluded: gene therapy is the future for medicine, although it will bring its risks and different views within society. Many races were studied to see a common trend, and within those trends and races, it makes sense for distrust in medicine, as scientists and doctors treat some patients as "guinea pigs." Although it was seen that with some of these injustices like with Henrietta Lacks and HeLa cells and her HeLa cells; that was still wrong and unethical to do. Especially without consent. Yes, it improved medicine, but at what cost? You can never be too safe with new medicines or gene therapy because procedures can fail, which is why consent is key.

Works Cited

Abraham, Allistair, and Rayne Rouce. “Supporting Gene Therapy Research and Access for Minority Populations.” American Society of Gene & Cell Therapy, 2025, www.asgct.org/publications/news/february-2021/black-history-month-sickle-cell. Accessed 26 June 2026.

Ahmadi, Ladan, and Kylie Murdock. “Five Inequities in Health Care.” Third Way, 2020, www.thirdway.org/report/five-inequities-in-health-care. Accessed 23 June 2026.

Bazargan, Mohsen, Sharon Cobb, and Shervin Assari. “Discrimination and Medical Mistrust in a Racially and Ethnically Diverse Sample of California Adults.” The Annals of Family Medicine, vol. 19, no. 1, 2021, pp. 4–15. DOI, doi:10.1370/afm.2632. Accessed 24 June 2026.

Beasley, Stephanie. “PhRMA, Minority Health Group Launch Effort to Diversify Clinical Trials.” FDA Week, vol. 20, no. 11, 2014, p. 16. JSTOR, www.jstor.org/stable/26692916. Accessed 27 June 2026.

Brandt, Allan M. “Racism and Research: The Case of the Tuskegee Syphilis Study.” The Hastings Center Report, vol. 8, no. 6, 1978, pp. 21–29. JSTOR, doi:10.2307/3561468. Accessed 23 June 2026.

Carpio, Myla Vicenti. “The Lost Generation: American Indian Women and Sterilization Abuse.” Social Justice, vol. 31, no. 4 (98), 2004, pp. 40–53. JSTOR, www.jstor.org/stable/29768273. Accessed 27 June 2026.

Centers for Disease Control and Prevention. “Health Disparities among LGBTQ Youth.” Adolescent and School Health, 6 Dec. 2024, www.cdc.gov/healthy-youth/lgbtq-youth/health-disparities-among-lgbtq-youth.html. Accessed 24 June 2026.

Harvard T.H. Chan School of Public Health. “Marginalized Communities Likely to Distrust Healthcare System, Polling Finds.” Harvard T.H. Chan School of Public Health, 7 Feb. 2023, hsph.harvard.edu/news/marginalized-communities-likely-to-distrust-healthcare-system-polling-finds/. Accessed 23 June 2026.

Office of Science Policy. “HeLa Cells: A Lasting Contribution to Biomedical Research.” National Institutes of Health, Oct. 2022, osp.od.nih.gov/hela-cells/. Accessed 25 June 2026.

KFF. “How History Has Shaped Racial and Ethnic Health Disparities: A Timeline of Policies and Events.” KFF, 2023, www.kff.org/how-history-has-shaped-racial-and-ethnic-health-disparities-a-timeline-of-policies-and-events/. Accessed 23 June 2026.

Lampe, Nik M., et al. “Health Disparities among Lesbian, Gay, Bisexual, Transgender, and Queer Older Adults: A Structural Competency Approach.” The International Journal of Aging and Human Development, vol. 98, no. 1, 2023, p. 009141502311718. SAGE Journals, doi:10.1177/00914150231171838. Accessed 24 June 2026.

Ndugga, Nambi, et al. “Racial and Ethnic Disparities in Access to Medical Advancements and Technologies.” KFF, 22 Feb. 2024, www.kff.org/racial-equity-and-health-policy/issue-brief/racial-and-ethnic-disparities-in-access-to-medical-advancements-and-technologies/. Accessed 23 June 2026.

National Coalition for LGBTQ+ Health. “Second State of LGBTQ Health™ National Survey.” HealthLGBTQ, healthlgbtq.org/stateof/lgbtqhealth/. Accessed 24 June 2026.

Saulsberry, Kalyn, and Sharon F. Terry. “The Need to Build Trust: A Perspective on Disparities in Genetic Testing.” Genetic Testing and Molecular Biomarkers, vol. 17, no. 9, 2013, pp. 647–648. DOI, doi:10.1089/gtmb.2013.1548. Accessed 23 June 2026.

Shahin, Wejdan, et al. “The Impact of Personal and Cultural Beliefs on Medication Adherence of Patients with Chronic Illnesses: A Systematic Review.” Patient Preference and Adherence, vol. 13, 2019, pp. 1019–1035. DOI, doi:10.2147/PPA.S212046. Accessed 24 June 2026.

Subica, Andrew M. “CRISPR in Public Health: The Health Equity Implications and Role of Community in Gene-Editing Research and Applications.” American Journal of Public Health, vol. 113, no. 8, 2023, pp. 874–882. doi:10.2105/AJPH.2023.307315. Accessed 24 June 2026.

Thompson, C., Robert A. Morgan, and R. Michael Blaese. “Gene Therapy.” BMJ: British Medical Journal, vol. 319, no. 7220, 1999, p. 1310. JSTOR, www.jstor.org/stable/25186364. Accessed 23 June 2026.

Yearby, Ruqaiijah, et al. “Structural Racism in Historical and Modern US Health Care Policy.” Health Affairs, vol. 41, no. 2, 2022, pp. 187–194. DOI, doi:10.1377/hlthaff.2021.01466. Accessed 23 June 2026.

Zúñiga, Nieves, et al. Does More Transparency Improve Accountability? Transparency International, 2018. JSTOR, www.jstor.org/stable/resrep20498. Accessed 27 June 2026.