

THE UNDERGRADUATE JOURNAL OF PHILOSOPHY

A Priori



BROWN UNIVERSITY
DEPARTMENT OF PHILOSOPHY

VOLUME IX
2026

A Priori

Brown University Undergraduate Journal of Philosophy

Volume IX



Brown University
Department of Philosophy
2026

Masthead

EDITORS-IN-CHIEF

Peter Haynes
Steve Nam
Austin Smart

DEVELOPMENTAL

EDITORS

Mansie Bennett
Romi Bhatia
Atiksh Gupta
Edward Ho
Maria Kim
Elena Ko
Jeffrey Kunzweiler
Charles Lebwohl
Charles Luke
Jason Mao
Eric Wu
Ziyi Ye

LINE EDITORS

Sophia Barki
Mansie Bennett
Avery Kaak
Lily Luby
Charles Luke
Jason Mao
Henry Roden
Zachary Woltanski

PRODUCTION

Bryanna Gonzalez
Elena Ko

COPY EDITORS

Bryanna Gonzalez
Maria Kim
Elena Ko
Jeffrey Kunzweiler
Gavin Li
Jason Mao
Daniel Mun
Henry Roden
Zachary Woltanski

READERS

Sophia Barki
Simon Beil
Mansie Bennett
Gabriel Gonzalez
Atiksh Gupta
Edward Ho
Avery Kaak
Maria Kim
Jeffrey Kunzweiler
Gavin Li
Charles Luke
Jason Mao
Eduardo Michelsen
Henry Roden
Eric Wu
Ziyi Ye

WEB DEVELOPMENT

Eric Wu

SOCIAL MEDIA

Emerson Rhodes

Letter from the Editors

Dear Readers,

We are excited to present Volume IX of *A Priori*: The Brown University Undergraduate Journal of Philosophy. Since its founding as the university's first philosophy journal, *A Priori* has aimed to present the finest essays in philosophy written by undergraduates from any institution. The papers in a volume of *A Priori* are committed to asking questions which can reveal the uncertainties behind our most confident attitudes and beliefs. In Volume IX, the papers inquire into topics spanning politics, morality, technology, history, and more. We hope the reader finds in Volume IX an opportunity to refresh their perspectives on issues that remain ever-present in the world.

The publication of this volume would not have been possible without the help of many friends of the journal. We wish to thank the Brown University Philosophy Department and the Philosophy Departmental Undergraduate Group for their financial support. Additionally, we are grateful to Katherine Scanga and Kalum Alldredge for their administrative assistance, especially in our communications with other undergraduate institutions. Most of all, we would like to thank every student on the *A Priori* team for their dedication and hard work.

This volume of *A Priori* is dedicated to Mukhammad Aziz Umurzokov and Ella Cook. May their memory continue to be a blessing to friends, family, and the Brown community.

Editors-in-Chief

Peter Haynes,

Steve Nam,

Austin Smart

Providence, Rhode Island, May 2026

Table of Contents

Hazing the Other: The “Brotherhood” We Never Rushed Lydia Jin	1
The Question Concerning Large-Language Models Callum Yeaman	13
bell hooks: Abuse and Imperfect Love Aidan Zhou	26
Interpellation in the Colonial Context: Louis Althusser’s Theory of Interpellation and Frantz Fanon’s Account of the Race-Marked Colonial Person Violeta Lopez Molina	32
Psychological Influences on Deceptionist Lies Francesca Stamati	41
The Significance of the Fleeting: Happiness and Mortality in Augustine’s <i>De Trinitate</i> Asal Arefi	47
Tainted Morality: Systemic Oppression and Moral Responsibility Madelyn Stanley	58
A Challenge to the Moral Argument: The Problem of Animal Suffering Adrian Haagen	69
Aesthetics as a Humanism Christina Pan	81
Bare Life in the Technopoly: The Potential for Democracy Under Surveillance Capitalism in the Age of Artificial Intelligence Rosalind Beattie	93

Hazing the Other: The “Brotherhood” We Never Rushed

Lydia Jin

Abstract: This paper argues that fraternity hazing functions as a microcosm of patriarchal oppression, revealing the mechanisms through which systems of domination reproduce themselves. Beginning with an analysis of institutional definitions of hazing, I demonstrate that their inadequacy reflects oppression’s inherent resistance to localization as discrete, punishable acts. Drawing on Arnold van Gennep’s concept of liminality and feminist critiques of consent, particularly Catharine MacKinnon’s analysis of coerced consent under conditions of power imbalance, I establish that hazing operates through a coercive social contract in which subordinates must consent to violation as the price of belonging. Employing Dianne Taylor’s Foucaultian analysis of humiliation and Sanjay Palshikar’s conception of humiliation as communication across power boundaries, I argue that “New Member Education” teaches pledges that subordinate members are unworthy of freedom, which is the same lesson women receive through patriarchal socialization. The paper traces structural parallels between the pledge’s predicament and women’s navigation of marriage, workplace harassment, and normative femininity: in each case, subjects face a double bind of submitting to degradation or forfeiting access to economic and social power. Engaging Åsa Burman’s work on structural power, I examine how intersecting oppressions of race, class, and gender compound within Greek life. I conclude that anti-hazing policies necessarily fail because they treat structural oppression as isolated incidents. Fraternities serve as training grounds where young men learn the contractual logic of domination they will later deploy in broader society. Recognizing hazing as structural rather than incidental demands that reform efforts address the exclusionary foundations of Greek life, not merely the symptoms of its hierarchy.

Since 1970, there has been at least one hazing-related death on a college campus each year.¹ In a case against two Florida A&M University students who beat Kappa Alpha Psi initiate Marcus Jones to the point of needing surgery, Honorable Judge Kathleen Dekker is quoted

¹ Elizabeth Allan and Mary Madden, *National Study of Student Hazing, Hazing in View: Students at Risk*, 2006 – 2008.

as saying: “I want to save the victims who will quietly go along because they want to belong.”² We gawk at headlines when initiation rituals lead to hospitalizations or death. We think of the victim’s families, and how they must feel having sent their child to college, only for them to die at the hands of careless college students with an invincibility complex on a power trip. It is even more disorienting when their deaths are ruled accidental, despite clear evidence that something more sinister was at play.

The concept of hazing is much older than college fraternities and dates back to Plato’s Academy, when boys were observed practicing pennialism, a system of mild oppression and torment practiced upon first-year students. The military is known for hazing, as well as many athletic and social groups. Although less likely to adopt traditional hazing practices, sororities cannot be exempt from the discussion either. Defenders of hazing will scoff at its adversaries, alleging that critics are simply bitter for not gaining acceptance into a club exclusive enough to be hazed. They insist that hazing promotes stronger bonds between initiates—a necessary evil for the pledges and an irreplaceable part of the fun for senior members of the community.³ The cycle of abuse seems impossible to break and as the practice lives on, hazers gain more entitlement to their actions through aging traditions.

Hazing may seem like an isolated case of misplaced power and torment, but it serves as a microcosm and therefore a perfect case study of large-scale oppression. This paper focuses on predominantly White fraternities and their relationship to patriarchal power, though the analysis illuminates other oppressive structures. Hazing relies on a coerced rite of passage that weaponizes humiliation, strips bodily autonomy, and forces participants to prove their worth through self-sacrificial submission. College fraternities serve as the training ground where young men learn and internalize the contractual logic of oppression. Patterns of behavior that resemble hazing appear in many contexts both covertly and overtly, suggesting that urgent attention should be paid to its proliferation and harmful effects. Addressing the problem of oppression requires careful analysis of its concentrated, smaller-scale forms. Fraternities teach young men that belonging is earned through submission, that past suffering entitles future cruelty, and that tradition renders these lessons uncontestable. I aim to examine how they teach the logic of the patriarchy itself.

The purpose of this paper is in no way meant to minimize the harm the victims of hazing face in its localized form. I instead wish to highlight its prevalence in more covert contexts and show how it directly mirrors the oppression we hope to eliminate. Hazing in Greek life exploits power differences and reinforces them to perpetuate the cycle of abuse. Fraternities and sororities are historically exclusionary institutions, which calls into question the value of their traditions and practices. The selection process for inclusion into the group is entirely based on perceptions of social power, which provides a clear gateway for oppressive behaviors such as hazing. Members are taught to accept, normalize, and perpetuate the oppression of subordinated groups.

2 “Frat Brothers Get 2 Years for Assault,” *CBS News*, January 30, 2007, <https://www.cbsnews.com/news/frat-brothers-get-2-years-for-assault/>.

3 It is interesting to note that Dartmouth College states in their Office of Greek life Handbook: “The term pledge is no longer used as it has the connotation of servitude or obedience.” Yet, due to the fact that new members do serve and obey as part of hazing, I deem it appropriate to use in the context of this paper. Hazers often give tasks to hazingees, from silly (e.g. “send brother X a weather report every day” or “serve brother X dinner at the dining hall”) to serious (e.g. “eat as many burritos until you vomit on your brothers” or “endure this hot poker brand on your skin”). The fact that hazingees comply is sufficient to show that they are serving and obeying their ‘pledgemasters.’

1. What Is Hazing? The Problem of Definition

1.1 Historical Context and Rites of Passage

Before we can understand hazing as a microcosm of oppression, we must confront the reality that the current definitions of hazing do not capture how hazing persists in reality. This definitional failure reveals something essential about the nature of oppression itself.

Hazing symbolizes a membership change from newcomer and outsider to member and teammate through a process akin to a rite of passage. The term ‘rite of passage’ was first coined by Arnold van Gennep, who describes a process of “separation, transition, and incorporation” that is observed by many cultures, often in religious and spiritual contexts, to mark important transition periods in one’s life.⁴ According to van Gennep, the function of a rite of passage is to signify a change in social standing, making it easy to identify rights and obligations between various members within a society. Repeated over generations, it helps maintain the norms and social hierarchies within the society. Van Gennep’s groundbreaking concept is the stage of liminality: a period during which a member is “betwixt and between” statuses. The parallels between van Gennep’s framework and fraternity hazing as a rite of passage is obvious. The pledgship period is when potential new members (PNMs) have accepted a bid, but have not yet been formally initiated into the group, despite their official membership recognized by the college or institution that the fraternity is affiliated with. This clearly puts them in van Gennep’s liminal stage, between the statuses of outsider and member.

1.2 Institutional Definitions and Their Inadequacies

It may be tempting to assume that we have now successfully determined a time frame within which hazing can occur. Yet, upon careful reading of the definitions provided by hazing policies in institutional handbooks, we see that this assumption is unfounded. While public conceptions of hazing indicate that it occurs during the liminal pledgship period, official definitions of hazing do not provide any time frame requirements.

The Department of Defense in 1997 defined hazing as follows:

Hazing is defined as any conduct whereby a military member or members, regardless of service or rank, without proper authority cause another military member or members, regardless of service or rank, to suffer or be exposed to any activity which is cruel, abusive, humiliating, oppressive, demeaning, or harmful. Soliciting or coercing another to perpetrate any such activity is also considered hazing. Hazing need not involve physical contact among or between military members; it can be verbal or psychological in nature. Actual or implied consent to acts of hazing does not limit the culpability of the perpetrator.⁵

Here, the Department of Defense has not provided any guidelines for when hazing might occur, nor for who might inflict it on whom. They do not even make clear whether there is a meaningful difference between an act of hazing and a petty insult. It would be senseless to consider every uncontested verbal insult as hazing, though this could plausibly fit within the definition.

4 Arnold Van Gennep, *Les Rites de Passage* (Librairie Félix Alcan, 1909).

5 Kirsten M. Keller et al. “Defining Hazing,” in *Hazing in the U.S. Armed Forces: Recommendations for Hazing Prevention Policy and Practice* (RAND Corporation, 2015), 5-6, <http://www.jstor.org/stable/10.7249/j.ctt19jcgpb.8>.

Turning back to college fraternities, where hazing gets its infamous reputation in popular culture, we hope for neater boundaries. Dartmouth College builds on New Hampshire’s statutory definition and considers an act to be hazing if it is a “condition of initiation into, admission into, continued membership in or association with any organization.”⁶ Once again, we have no reason to assume that hazing must occur in a specific liminal period. Dartmouth’s policy similarly indicates that any member at any time could be hazed, now with the added condition that a victim’s lack of conformity should indicate a risk of losing membership in the group.

The definition implies that the “risk of losing membership” is associated with the risk of losing official membership, but I posit that being socially isolated, feeling uninvited, or othered by the group is effectively equal to losing official membership status. This should not be controversial, since van Gennep’s liminal stage and the pledgship period shows that there exists an in-between stage where members can be in a state other than part of the in-group, even if their official affiliation is with the group. Therefore, any future mention of membership within this paper should be taken to mean both official membership and unofficial social inclusion.

What could be considered necessary for “continued membership in or association with” the organization is largely up for interpretation. Organizations are always at liberty to state post facto that nobody’s membership was truly at risk and that participation in any activity was entirely optional. Even in the case where someone says ‘do this or you are out’ or ‘you are being hazed,’ it can be unclear whether these declarations are meant to be taken seriously or ironically.

In July 2024, Dartmouth student Won Jang drowned in the Connecticut River due to hazing. One fraternity member even wrote in a group chat that night: “Hazing is off to a phenomenal start.”⁷ Still, the Hanover Police Department ruled Jang’s death an accident.⁸ When nothing, not even consequences as permanent and traumatic as death with explicit, written acknowledgement of hazing is strong enough evidence to satisfy legal standards for institutional accountability, we must confront our dire inadequacy in understanding the matter at hand. This legal failure reflects the conceptual inadequacy of treating hazing as a discrete act rather than a structural one.

In both the Department of Defense’s and in Dartmouth College’s definition of hazing, the range of activities that could constitute as part of a hazing act is extremely broad. What one considers “cruel, abusive, humiliating, oppressive, demeaning, or harmful” can vary. A simple insult may suffice if directed at a particularly sensitive individual. Yet an insult is not enough for a hazing conviction. So, what is enough?

How does hazing differ from bullying? Bullying, while also not without its ambiguities, can be more easily identified from the victim’s point of view. If person B claims to be bullied by person A, one need only to observe A’s behavior to verify B’s case.⁹ B’s acknowledgement

6 For a full definition of Dartmouth’s hazing policy, see Dartmouth College, “Hazing Policy,” *Policies*, June 20, 2025, <https://policies.dartmouth.edu/policy/hazing-policy>. In this section, I am drawing on only a small section of the policy to illustrate that its timeframe boundaries are vague. The full policy still does not provide adequate definition, the reasons for which will become clear in further sections of this paper.

7 Susan Zalkind, “Greek Tragedy: A Drowning at Dartmouth College,” *Boston Magazine*, April 7, 2025, <https://www.bostonmagazine.com/news/2025/04/06/dartmouth-college-fraternity-culture-student-death/>.

8 The Dartmouth Senior Staff, “Won Jang’s Death Ruled Accidental by Police,” *The Dartmouth*, September 19, 2024, <https://www.thedartmouth.com/article/2024/09/won-jangs-death-ruled-accidental-by-police>.

9 By “observe,” I mean in a God’s-eye-view sense. Of course, A could easily conceal their bullying to human onlookers. In knowing A’s true behavior towards B, we can verify that bullying is taking place. A’s bullying of B is empirically verifiable, even if it may not always be easy to do.

that they are being harmed is part of the evidence that they are being harmed. Bullying is non-consensual. Hazing, on the other hand, must be consensual, even if that consent is not explicit. In the next section, I expand on how consent is required for hazing. Can someone be faulted for bullying if their victim has consented to it? If someone offered to pinch me every day and I let them, is the pincher morally reprehensible?

2. The Mechanics of Hazing as Oppression

2.1 Power Imbalances and Consent

The inadequacy of these definitions points us towards seeing hazing as a power structure sustained by coerced consent. To understand this, I examine how hazing resembles rape, another act cloaked in the language of consent.

Within feminist literature, conversations about consent have focused on physical freedom or resistance of women in and out of sexual contexts. Rape was considered an act of violence that restricted the physical freedom of the victim. Moreover, women were physically constrained in moving through society in nonsexual contexts out of fear of rape. According to Susan Rae Peterson, “what is wrong with rape is primarily its restriction of the freedom of bodily movement for women.”¹⁰ Further, she says that “rape effectively keeps women in their places; indeed, because many women fear being raped, they remain much more stationary and sedentary than men.” This proposition goes further than the common law that defined rape, which stated that nonconsensual sex occurred when a woman had physically resisted to sex, or was physically constrained in a way that she could not resist sex.¹¹ Yet it still had to be acknowledged that nonconsensual sex is a violation even when it does not involve physical resistance. The focus then shifted to consent in general, not only as a demonstration of physical compliance or resistance.

Some feminists questioned the validity of consent in general within the patriarchal structures of society. One argument is that where there exist clear power imbalances, consent cannot be uncoerced, and thus cannot be granted. This is recognized by several instances of the law, such as in relationships between teachers and students, therapists and patients, and adults and children. In cases like these, the power dynamic “is such that consent might not easily be refused.”¹² Catherine MacKinnon famously took this notion to show that in a male-dominated society, sex between a man and a woman is always fraught with power imbalances, which makes it doubtful that a woman can autonomously consent to sex. Readers and many critics of MacKinnon misinterpret her words to mean that “all sex is rape.”¹³ While this particular controversy is not the focus of this paper, the point that mixing power and consent creates muddy boundaries is salient.

The definitions of hazing that I have presented have stated that any member, “regardless of service or rank” in the case of the military, may be both perpetrators and victims of

10 Susan Rae Peterson, “Coercion and Rape: The State as a Male Protection Racket,” in *Feminism and Philosophy*, ed. Mary Vetterling-Braggin, Frederick A. Elliston, and Jane English, Littlefield, Adams & Co., 1977, 366.

11 See, e.g., *Rusk v. State*, 406 A.2d 624 (Md. Ct. Spec. App. 1979), rev’d, 424 A.2d 720 (Md. 1981).

12 A teacher-student case; *State v. McKenzie-Adams*, 915 A.2d 822, 836 (Conn. 2007), overruled on other grounds by *State v. Payne*, 34 A.3d 370 (Conn. 2012), and a lieutenant colonel-private first-class case; *Loomis v. United States*, 68 Fed. Cl. 503, 519 (2005).

13 Catherine MacKinnon and Andrea Dworkin instead aimed to show that consent is not a stable enough foundation to ground rape theory in.

the crime. Several examples illustrate this point. Being coerced into hazing another member may itself constitute a form of hazing, as there is evidence to suggest that hazing others also produces both negative short- and long-term psychological effects.¹⁴ This indicates that not all hazers find pleasure in committing the act, but are doing so to conform to the group's norms and practices. A member who refuses to haze others may not be expelled from the group explicitly, but as they continue to refuse to conform to the norms associated with their membership, they may become socially isolated and effectively excluded. As I have already stated, effective exclusion from the group in this way is practically equal to having membership from the group explicitly revoked. Thus, their membership being at risk means that some hazers may be considered as being hazed themselves by others in their cohort.

While a member could haze another member of the same age, rank, or cohort, hazing necessarily involves a power imbalance which is easily overlooked by the wording given in standard definitions. This imbalance may not present itself through official titles or positions, but through implicit dynamics such as ones based in racial, sexual, or cultural norms. Furthermore, social structures between members of a cohort play important roles in providing indirect forms of power to certain members of the group. Åsa Burman's work in social ontology highlights how social structures can equip agents with positive or negative structural power.¹⁵ For fraternities and sororities, racial exclusion is baked into many of their traditional practices and attitudes. Thus, racial differences can effect structural power between members in the same cohort. Additionally, things like wealth, attractiveness, family background, college majors, or general social perception can significantly alter the power dynamics within a cohort to varying degrees. With this in mind, it is easy to see how those with significant social power, such as a wealthy, tall, White male athlete may dominate another member in the same cohort who is a non-wealthy, short, Asian NARP.¹⁶ The more powerful can coerce the less powerful to participate in a dangerous activity through fear that non-conformity will cause him to lose membership, explicitly or otherwise, in the group.

2.2 Intersecting Oppressions and Structural Power

These intersecting oppressions compound within Greek life in ways that require careful attention. A Black woman in a historically White sorority faces both the hazing structure and racist microaggressions that function as continuous, informal hazing. Her 'choice' to remain is even more constrained, as leaving means losing not just social power but one of few institutional paths to professional networks in a White-dominated economy. Greek organizations were founded on exclusionary principles of race, class, and gender. Historically Black Greek Letter Organizations arose precisely because Black students were excluded from White Greek life. The intersectional dimensions of Greek life oppression deserve fuller treatment than this paper's scope allows, but they must also be acknowledged as constitutive to how hazing functions.

2.3 Coerced Consent and the Contractual Structure

Hazing involves a form of voluntary submission. However, as in the case of rape, coerced consent is not consent. Part of what makes coerced consent in rape so insidious is that

14 Hank Nuwer, *Hazing: Destroying Young Lives*, Indiana University Press, 2018.

15 Åsa Burman, *Nonideal Social Ontology*, Oxford University Press, 2023, 221.

16 NARP: Dartmouth slang for Non-Athletic Regular Person. See <https://journeys.dartmouth.edu/folklorearchive/narp/>.

the victim is coerced into saying yes not to sex, but to rape. Similarly, in cases of hazing, where the victims must perform some activity that may or may not be directly harmful, they are saying yes not to the activity, but to being hazed. While rape victims can also be violated without any consent, coerced consent is present in all cases of hazing, which makes hazing difficult to identify. Hazing requires its victims to consent to being violated, which must be considered a violation in itself. This form of coerced consent is what gives hazing its character in the broadest sense. Defining it this way opens the doors to seeing its prevalence in society at large.

The contractual structure that hazing invokes means that hazing necessarily involves consent from its victim. Simply put, if a member does not wish to be hazed, they will exit the group. This could mean that they exit officially or unofficially, through being socially isolated or excluded from the in-group. By continuing their membership in the group, they are indicating their consent to be hazed. Though their true desire is not the activity of hazing, but inclusion in the group, since inclusion in the group requires being hazed, they have no choice but to consent. Thus, their consent is coerced. Michael Cholbi argues that hazing includes a contractual arrangement, and this "need not be explicitly stated in order for a quid pro quo arrangement to be in force."¹⁷ Cholbi notes that "a reasonable person would, in the context, recognize that her complying with the intended harm is a condition of her full inclusion in the group or organization."¹⁸ In other words, hazing is a coercive social contract where the price of full membership is consent to being harmed.

3. Humiliation as Education

3.1 The Nature of Humiliation

Having established hazing as oppression sustained by coerced consent across power differentials, we can now examine what this oppression teaches its victims.

Many confuse coerced consent in this context with some masochistic tendency—perhaps the victims actually enjoy being hazed, which is why they agree to it in the first place. Sanjay Palshikar excludes masochism from his discussion of humiliation, just as I will here.¹⁹ Since the majority of activities involved in hazing rituals subject victims either to immediate or long-term physical and psychological harms, and the fact that most victims would not voluntarily engage in them outside of the context of obtaining membership in the group, it is reasonable to assume that victims are not gaining pleasure from hazing rituals, and thus not engaging in masochism, but are being subjected to humiliation.

Dianna Taylor discusses humiliation as a disruption of the Foucaultian view of the self.²⁰ Michel Foucault focuses on the interconnectedness of a relationship with oneself and that of the external world. We cannot separate ourselves from others or their treatment of us; a person's sense of self is a constant hermeneutics across these relationships. An act of humiliation "threatens to completely foreclose those capacities" and "has the potential to fix the self in an abject relation in which it is displayed before itself merely as a negation, as not only

17 Michael J. Cholbi, "On Hazing," *Public Affairs Quarterly* 23, no. 2 (2009): 143–60, <http://www.jstor.org/stable/40441523>, 144.

18 Cholbi, "On Hazing," 145.

19 Sanjay Palshikar, "Understanding Humiliation," *Economic and Political Weekly* 40, no. 51 (2005): 5428–32, <http://www.jstor.org/stable/4417556>.

20 Dianna Taylor, "Humiliation as a Harm of Sexual Violence: Feminist versus Neoliberal Perspectives," *Hypatia* 33, no. 3 (2018): 434–50, <http://www.jstor.org/stable/45153703>.

unworthy but indeed incapable of freedom.”²¹ Palshikar asserts that humiliation is a form of communication across power boundaries and cannot, by definition, be one-sided.²² While it is debated whether the perpetrator of humiliation sees the victim as human or non-human, there is consensus that the victim sees themselves as humiliated.²³

3.2 “New Member Education”

Hazing is often colloquially clothed under the term ‘New Member Education,’ or NME. But what, exactly, are new members being educated on? Perhaps they are forced to recite facts about the organization,²⁴ or learn the names of their fellow pledges. Their ‘education’ may be seen as a way of bonding with other members of the group, particularly the cohort of victims; friendships and solidarity often mitigate the pain that insults wreak.²⁵ Yet diluting the harms through camaraderie also works as kindling for the fire, as hazers view the harm they have inflicted in the moment as less than it actually is. Moreover, it circumvents any confrontation of long-term harms to an individual, or ripple effect harms that these individuals may then inflict on future cohorts.

In return for acceptance into the group, pledges of a fraternity must first acknowledge their position as lowest in the pecking order, having no influence or agency within the domain of the organization. Combining Palshikar’s notion of humiliation as communication across power lines and Taylor’s analysis of humiliation and how it affects the self, I assert that the education received in NME is that *subordinate members of a society are not worthy of freedom*. By freedom, I mean the ability to define and interpret oneself within the context of others as one sees fit, as opposed to as someone in a higher position of power sees fit. Pledges recognize their subordinate status with the knowledge that retaliation results in being outcast from the society altogether. The decision not to retaliate, as Palshikar puts it, “means that the situation is not seen as confronting us with a momentous choice, it means that the history-making potential of ‘humiliation’ is not explored.”²⁶ As future generations enter the system, the “momentous choice” and “history-making potential” grows further and further from view, and eventually, it becomes accepted that ‘this is just how it’s done.’

3.3 *The Parallel to Women’s Socialization*

This education in subordination is not unique to fraternities. Women receive the same curriculum, administered through different institutions with different rituals, but with parallel underlying logic.

Women learn through countless mediums that we are not worthy of defining ourselves on our own terms, but only as men define us. We know that if we do not conform to the patriarchal standards of womanhood, then we cannot thrive in the society that has been built through teaching that same lesson, time and time again. As bell hooks argues in *The Will to Change*, patriarchy demands that women perform subordination as the price of economic survival and

21 Taylor, “Humiliation as a Harm of Sexual Violence,” 440.

22 Palshikar, “Understanding Humiliation,” 5431.

23 Taylor discusses these two views held by Avishai Margalit and Lisa Guenther. Margalit argues that the perpetrator sees the person as human, but treats them as non-human, while Guenther argues that the perpetrator sees the person as non-human, and treats them as such. Both agree that the victim sees themselves as humiliated.

24 As Eduardo Saverin does when joining a finals club at Harvard in *The Social Network* (2010).

25 Palshikar, “Understanding Humiliation,” 5429.

26 Palshikar, “Understanding Humiliation,” 5431

social belonging, which is precisely the contractual logic of hazing.²⁷ Women’s health research is chronically underfunded and women, and even female mice, were excluded from clinical trials until 1993.²⁸ Matters of our existence, as members of the low, is “beneath concern” for society.

4. Hazing as Microcosm of Patriarchal Oppression

4.1 *The Contractual Nature of Women’s Oppression*

In the patriarchy, women are coerced into a contractual arrangement in which they submit to oppressive norms and practices in exchange for economic, legal, and social security. The workplace ‘initiation’ women face includes enduring sexual harassment as ‘just how it is,’ performing emotional labor to prove they’re ‘team players,’ and tolerating interruptions and idea-theft to maintain their seat at the table. Like fraternity pledges, professional women face a choice: to submit to degradation or lose access to power. The #MeToo movement revealed how women across industries had ‘consented’ to harassment for decades—not because they wanted to, but because refusing meant career death. Harvey Weinstein’s assistants knew they were being humiliated; they also knew that objecting meant blacklisting.

Marriage historically functions as another contractual form of this arrangement. In her book, *The Case Against the Sexual Revolution*, Louise Perry writes that marriage “has historically been used as a vehicle for the control of women by men,” but that women should marry anyway because there is no better system.²⁹ Marriage functions as a vehicle for protection for women, particularly if they have children. Marriage has a “taming effect” on men, and “a society composed of tamed men is a better society to live in, for men, for women and for children.”³⁰ Women who refuse marriage face economic precarity in a system designed to punish non-compliance. I do not mean to say that every woman must enter a heterosexual marriage in order to thrive in society. However, there is strong reason to think that doing so comes with undeniable benefits, despite the institution of marriage being simultaneously oppressive. Other institutions demonstrate a similar contradiction, such as normative femininity itself. In exchange for protection and inclusion in greater society, women must conform to systems that fuel their own subordination.

These parallels aim to show the similarities between the oppression of women in the patriarchy and of hazing victims in college fraternities. These two structures are clearly related because their dynamics broadly function in the same ways. In both cases, continued access to status, security, and membership depends on the willingness to endure practices that reinforce one’s own subordination.

4.2 *Bonding Through Shared Trauma*

Defenders of hazing often insist that it brings members of the group closer emotionally. Psychologically, there is some basis for this claim. The severity-affiliation-attraction hypothesis suggests that the more suffering individuals endure for the purpose of an objective, the more they are forced to resolve the cognitive dissonance of why they are in that position, and the more value they place on that objective.³¹ In this case, the objective is being recognized as part

27 bell hooks, *The Will to Change: Men, Masculinity, and Love*, Washington Square Press, 2004.

28 Bridget Balch, “Why We Know so Little about Women’s Health,” *Association of American Medical Colleges*, March 26, 2024, <https://www.aamc.org/news/why-we-know-so-little-about-women-s-health>.

29 Louise Perry, *The Case against the Sexual Revolution*, Polity, 2022, 184.

30 Perry, *The Case against the Sexual Revolution*, 182.

31 Jason M. Silveira and Michael W. Hudson, “Hazing in the College Marching Band,” *Journal of Re-*

of the in-group. On this view, hazing promotes group bonding through shared trauma; pledges endure suffering together and later reinterpret that suffering as proof of loyalty and depth of connection to the group.³² Additionally, as mentioned earlier, friendship can lessen the pain of trauma. Thus, hazing victims form bonds with each other in order to weaken their suffering.

However, while bonds formed through suffering can be powerful, an attempt to forge these connections cannot be a reasonable justification to inflict harm. Native people do not thank colonizers, slaves do not thank enslavers, and women do not thank men for the strength or value of their respective communities, culture, and camaraderie. Instead, they find power in sharing their individual identities and who they become alongside their trauma, not because of it. As I have stated in previous writing, we take power away from the system when we choose to prioritize our individuality over who the system forces us to become.³³

Proponents might object that voluntary military service also involves hardship that bonds soldiers, yet we don't condemn basic training as hazing. The distinction lies in purpose, equality, and structural position. Military training's hardships serve external objectives such as combat readiness and are imposed equally on all recruits by the institution itself, not by those who recently endured the same treatment seeking to reassert dominance. Hazing serves only the internal objective of maintaining hierarchy and is imposed by those with arbitrary social power over those without it. It is oppression for oppression's sake, dressed up as tradition. Yet even this distinction should give us pause—military hazing does occur, and military culture trains men in domination and violence in ways feminist scholars have extensively documented.³⁴ The military-fraternity parallel may be closer than the objection suggests, both serving as masculinity-training institutions that prepare men for lives of dominance.

4.3 Why This Matters: From Fraternity to Society

The fraternity system trains young men not just to accept hierarchy, but to glorify it, finding pleasure and meaning in their power over others, and to view submission from others as their due. These men then carry this training into boardrooms, courtrooms, and bedrooms. They have learned that love, loyalty, and belonging are proven through submission to humiliation. They have learned that their own past suffering entitles them to inflict future suffering. They have learned that tradition justifies cruelty, that consent can be manufactured through coercion, and that questioning the system means exclusion from power. In short, they have learned to be patriarchs. And women, whether or not we ever rushed a sorority, have been pledging to their brotherhood our entire lives.

5. The Moral and Practical Failure

5.1 Why Anti-Hazing Policies Fail

Hazing has been outlawed in 44 of the 50 states, and in the District of Columbia.³⁵

search in Music Education 63, no. 1 (2015): 5–27, <http://www.jstor.org/stable/43900277>.

32 Keller et al. suggest that humiliation may also cause negative attitudes towards the group. The data is inconclusive, so my aim in this section is to make a case against those who believe that humiliation is a necessary component of group bonding.

33 Lydia Jin, "Jin: The Price of Our Community: Paying with Your Life," *The Dartmouth*, May 6, 2025, <https://www.thedartmouth.com/article/2025/05/jin-the-price-of-our-community>.

34 Maya Eichler, "Militarized Masculinities in International Relations," *The Brown Journal of World Affairs* 21, no. 1 (2014): 81–93, <http://www.jstor.org/stable/24591032>.

35 "States with Anti-Hazing Laws | StopHazing | Hazing Prevention Resource," *StopHazing*, n.d., <https://>

On the surface, we see a moral obligation to protect our youth from the harms of hazing. Iris Marion Young's "Five Faces of Oppression" helps clarify what exactly makes hazing morally wrong. Hazing combines exploitation (extracting emotional/physical labor from subordinates), marginalization (threatening exclusion), powerlessness (stripping agency), cultural imperialism (defining pledges' reality for them), and violence (psychological and physical harm).³⁶ Hazing is exceptional in being corrupt from the outset, since its constitutive aim is to compel those hazed into a subordinate posture incompatible with their status as the rational coequal of those who haze them.

The reason definitions of hazing are so broad and thus effectively useless is that they attempt to localize a problem that cannot be localized. A policy that outlaws hazing is really reaching for something deeper that has been bound with human suffering for as long as society has existed: oppression itself. If we are to address the issue of hazing, we must eliminate the social contexts that facilitate it. We cannot prevent hazing through attempting to define it as distinct from what it truly is. Fraternities and sororities inherently exclude others based on social power, and where there is exclusion based on social power, there will always be the oppression of less powerful groups.

Anti-hazing policies fail because they treat hazing as a discrete act rather than a systemic structure. Universities prohibit 'forcing someone to consume alcohol' but cannot prohibit the social architecture that makes refusal costly. They cannot ban the power differentials that make interactions between members coercive. The problem isn't the specific act; it's the system that makes the act rational for both perpetrator and victim. Until we address exclusionary institutions built on social power differentiation, we are merely treating the symptoms instead of curing the disease.

6. Conclusion

Hazing in fraternities is a concentrated display of the same logic that structures patriarchy and other large-scale oppressions: coerced consent under conditions of unequal power, ritualized humiliation as "education," contractual arrangements, and a double bind that forces the oppressed to either submit or be cast out. By training members to accept humiliation as the price of belonging, hazing rehearses and reproduces the very patterns of domination that shape our broader social world. The definitional uncertainty of hazing comes from oppression's inherent resistance to clean boundaries precisely because of its foundations in social relations rather than isolated incidents.

This analysis has focused on predominantly White fraternities where young men learn to be patriarchs, but the implications extend further. The contractual logic of "consent to subordination or be excluded" operates implicitly in any structure where there exists a power hierarchy. Recognizing hazing as structural means recognizing that reform will never be sufficient. We cannot make Greek life safe while preserving its exclusionary foundations any more than we can make patriarchy safe while preserving male dominance. The question is not how to define hazing more precisely, but whether we are willing to abolish the institutions that require its logic to function.

stophazing.org/policy/state-laws/.

36 Iris Marion Young, *Justice and the Politics of Difference*, Princeton University Press, 1990.

Works Cited

- Allan, Elizabeth, and Mary Madden. *National Study of Student Hazing: Hazing in View: Students at Risk*. 2006–2008.
- Balch, Bridget. “Why We Know So Little about Women’s Health.” *Association of American Medical Colleges*. March 26, 2024. <https://www.aamc.org/news/why-we-know-so-little-about-women-s-health>.
- Burman, Åsa. *Nonideal Social Ontology*. Oxford University Press, 2023.
- Cholbi, Michael J. “On Hazing.” *Public Affairs Quarterly* 23, no. 2 (2009): 143–60. <https://www.jstor.org/stable/40441523>.
- Dartmouth College. “Hazing Policy.” *Policies*. June 20, 2025. <https://policies.dartmouth.edu/policy/hazing-policy>.
- Eichler, Maya. “Militarized Masculinities in International Relations.” *The Brown Journal of World Affairs* 21, no. 1 (2014): 81–93. <https://www.jstor.org/stable/24591032>.
- “Frat Brothers Get 2 Years for Assault.” *CBS News*. January 30, 2007. <https://www.cbsnews.com/news/frat-brothers-get-2-years-for-assault/>.
- hooks, bell. *The Will to Change: Men, Masculinity, and Love*. Washington Square Press, 2004.
- Jin, Lydia. “Jin: The Price of Our Community: Paying with Your Life.” *The Dartmouth*. May 6, 2025. <https://www.thedartmouth.com/article/2025/05/jin-the-price-of-our-community>.
- Keller, Kirsten M., Miriam Matthews, Kimberly Curry Hall, William Marcellino, Jacqueline A. Mauro, and Nelson Lim. “Defining Hazing.” In *Hazing in the U.S. Armed Forces: Recommendations for Hazing Prevention Policy and Practice*, 5–22. RAND Corporation, 2015. <https://www.jstor.org/stable/10.7249/j.ctt19jcgpb.8>.
- Nuwer, Hank. *Hazing: Destroying Young Lives*. Indiana University Press, 2018.
- Palshikar, Sanjay. “Understanding Humiliation.” *Economic and Political Weekly* 40, no. 51 (2005): 5428–32. <https://www.jstor.org/stable/4417556>.
- Perry, Louise. *The Case against the Sexual Revolution*. Polity, 2022.
- Peterson, Susan Rae. “Coercion and Rape: The State as a Male Protection Racket.” In *Feminism and Philosophy*, edited by Mary Vetterling-Braggin, Frederick A. Elliston, and Jane English. Littlefield, Adams & Co., 1977.
- Silveira, Jason M., and Michael W. Hudson. “Hazing in the College Marching Band.” *Journal of Research in Music Education* 63, no. 1 (2015): 5–27. <https://www.jstor.org/stable/43900277>.
- “States with Anti-Hazing Laws | StopHazing | Hazing Prevention Resource.” *StopHazing*. n.d. <https://stophazing.org/policy/state-laws/>.
- Taylor, Dianna. “Humiliation as a Harm of Sexual Violence: Feminist versus Neoliberal Perspectives.” *Hypatia* 33, no. 3 (2018): 434–50. <https://www.jstor.org/stable/45153703>.
- The Dartmouth Senior Staff. “Won Jang’s Death Ruled Accidental by Police.” *The Dartmouth*, September 19, 2024. <https://www.thedartmouth.com/article/2024/09/won-jangs-death-ruled-accidental-by-police>.
- Van Gennep, Arnold. *Les Rites de Passage*. Librairie Félix Alcan, 1909.
- Young, Iris Marion. *Justice and the Politics of Difference*. Princeton University Press, 1990.
- Zalkind, Susan. “Greek Tragedy: A Drowning at Dartmouth College.” *Boston Magazine*. April 7, 2025. <https://www.bostonmagazine.com/news/2025/04/06/dartmouth-college-fraternity-culture-student-death/>.

The Question Concerning Large-Language Models

Callum Yeaman

Abstract: In 1954, Heidegger’s “The Question Concerning Technology” warned that technology’s true danger emerges when we mistakenly view it as something “neutral,” urging rigorous questioning of our relation to emerging technologies. Today, nearly 75 years later, technological development has spiraled to heights likely inconceivable in Heidegger’s time—most recently with the arrival of artificial intelligence. This paper revisits Heidegger’s seminal work to undertake a phenomenological analysis of large-language models (LLMs).

I begin in Section II by expounding Heidegger’s phenomenological method and unpacking his key terminology, clarifying what it means to question technology’s essence. In Section III, I explore Heidegger’s claim that technology’s essence is a “realm of revealing,” adopting this as my guiding framework for investigating LLMs. Section IV argues that LLMs, as a realm of revealing, disclose our relationship to language and understanding—what Heidegger explores through “discourse.” In Section V, I unpack Heidegger’s theses on authentic and inauthentic discourse, explaining why the degree to which discourse is inauthentic correlates to how thematic it becomes. Finally, in Section VI, my investigation culminates in the conclusion that the essence of LLMs is idle talk. My central aim is not merely to argue that LLMs exemplify inauthentic discourse, but to suggest that their proliferation reveals—and is enabled by—a “flattened” understanding of language, one that reflects the pervasive influence of what Heidegger calls the they.

I. Introduction

In 1954, Heidegger’s “The Question Concerning Technology” warned humanity to be wary of technology’s profound grip on modern existence. The true danger of technology arises, he argued, when we make the grave mistake of viewing it as something “neutral.”¹ Amidst a time of rapid change, Heidegger urged us to recognize the necessity of rigorously questioning our relation to emerging technologies.

Today, nearly 75 years after Heidegger published his seminal essay, technological

¹ Martin Heidegger, “The Question Concerning Technology,” in *The Question Concerning Technology, and Other Essays*, trans. William Lovitt (Harper & Row, 1977), 1.

development has spiraled to heights that were likely inconceivable in his time. In the age of artificial intelligence, it's possible our relationship to technology has never been so mysterious. In this paper, I revisit Heidegger's renowned essay in the epoch of artificial intelligence. This paper seeks to apply Heidegger's phenomenological mode of inquiry to explore the essence of large-language models (hereafter, LLMs).

Before I approach "The Question Concerning Technology," I begin Section II by briefly expounding Heidegger's phenomenological method and unpacking some of his key terminology, as well as explaining what it means to "question" technologies' essence. Then, in Section III, I unpack Heidegger's nebulous claim that technology's essence is a "realm of revealing." I adopt this definition as my guiding framework as I investigate the essence of LLMs.

In Section IV I argue that, as a realm of revealing, LLMs disclose our relationship to language and understanding—what Heidegger explores through the term "discourse." My leading contention is that an analysis of the architecture of LLMs reveals the prescience of Heidegger's writings on *authentic* and *inauthentic* discourse and, in particular, the pervasiveness of what he calls *idle talk*. My central aim is not just to argue that LLMs exemplify *inauthentic discourse*, but to suggest that their proliferation reveals—and is enabled by—a "flattened" understanding of language. In Section V, I unpack Heidegger's central theses on *authentic* and *inauthentic* discourse, and explain why he claims the degree to which discourse is *inauthentic* is correlated to how *thematic* it is. Finally, in Section VI, I argue that the architecture of LLMs reflects the prevalence of highly *thematic*, *inauthentic* forms of discourse today. LLMs reveal the pervasive influence of *idle talk*: the way the "flattening" of language engenders the domination of what Heidegger calls *the they*.

II. On Questioning: Heidegger's Phenomenological Method

"Questioning builds a way,"² mysteriously begins Heidegger's seminal paper, "The Question Concerning Technology." I seek to question the essence of LLMs. But what does it mean to *question* the essence of something? According to Heidegger, the ultimate purpose of questioning something's essence is to expose our relation to it.³ This is the case because the broader aim of Heidegger's analytic method is to elucidate the meaning of a word he claims has been overlooked throughout the history of philosophy: *being*.⁴ In order to achieve this, Heidegger often relies upon commonly used words and develops his own unconventional definitions for them (e.g., *discourse*, *assertion*, *authentic*). I will indicate such terms by italicizing them. Heidegger believes that an inquiry into something's essence must always begin from an examination of our own sort of being: what he refers to as *Dasein*⁵ (although often left untranslated, this most literally translates to *being-here* or *openness*).⁶ Heidegger views it as futile to contemplate the essence of entities in the world by examining them over and against a subject—because the being of *Dasein* and the being of entities in the world can't successfully be separated in philosophical inquiry.⁷ Therefore, he insists phenomenological inquiry must begin by investigating the joint structure of *being-in-the-world*.

2 Heidegger, "The Question Concerning Technology," 1.

3 Daniel O. Dahlstrom, *The Heidegger Dictionary* (Bloomsbury Academic, 2013).

4 Martin Heidegger, *Being and Time*, trans. John Macquarrie and Edward Robinson (Harper & Row, 1962).

5 Heidegger, *Being and Time*.

6 Dahlstrom, *The Heidegger Dictionary*.

7 Heidegger, *Being and Time*.

An inquiry into *being-in-the-world* could not begin by investigating the material composition of particular entities, a mode of inquiry Heidegger refers to as *ontic*.⁸ An *ontic* inquiry into a butter knife would investigate the knife's build, measurement, and properties. No more can an inquiry into *being-in-the-world* proceed through an investigation of the *being* of those entities, which he defines as *ontological* inquiry.⁹ *Ontological* inquiry investigates the universal properties of butter knives—like, as Heidegger will argue, that a butter knife is something that cuts. Nor can a proper examination of *being-in-the-world* begin by analyzing facts about *Dasein*'s current circumstances, which are *existentiell* properties, in Heidegger's language.¹⁰ The object of *existentiell* inquiry might be to think of how and why you most recently used a butter knife. According to Heidegger, phenomenological inquiry must begin by expounding the *ontological* structures of *Dasein's existence*, which he calls *existential* properties.¹¹ *Existential* properties are what Heidegger refers to as the "internal enablement" of either an entity's *ontic* manifestation or the emergence of an *existentiell* fact.¹² This means that the *existential* structures of *Dasein's existence* are what make it possible for something (either *ontic* or *existentiell*) to manifest externally in the way that it does. For example, the existence of a butter knife is *existentially* enabled by the structures of *Dasein's existence* which render it possible for us to conceive of cutting something (most importantly of which, as I will explain in Section IV, is our pre-reflective relationship to time). As Heidegger explains in his later lectures, "On the Essence of Truth," when questioning the essence of something—whether technology or a butter knife—it's always necessary to begin by investigating its *existential* enablement.¹³ By exploring the *existential* enablement of LLMs, this paper will investigate what technology and LLMs reveal about *being-in-the-world*.

III. The Question Concerning Technology

Despite the relentless vicissitudes of technological development, Heidegger doesn't see the "essence" of technology as something indeterminate. Undoubtedly, technology is an amorphous, continually evolving force, which might make its essence seem impossible to pin down. How is it possible to reach conclusions about the essence of technology that hold across thousands of years of human evolution? According to Heidegger, such misgivings stem from a fundamental misunderstanding. For Heidegger, the essence of technology is "by no means anything technological."¹⁴ In other words, when inquiring about the essence of technology, it's pernicious to focus on qualities or properties of particular artifacts, because we shouldn't be questioning the essence of technology from an *ontic* standpoint in the first place.

Heidegger traces every purported attempt to capture the essence of technology back to Aristotle's fourfold causal definition. Aristotle claims that behind any *techné* (craft) are four causes: *i*) the material cause: the materials the craft is made out of *ii*) the formal cause: the general "shape" or form of the artifact *iii*) the maker cause: the builder who produces it

8 Heidegger, *Being and Time*.

9 Heidegger, *Being and Time*.

10 Heidegger, *Being and Time*.

11 Heidegger, *Being and Time*.

12 Heidegger, *Being and Time*.

13 Daniel O. Dahlstrom, "Im-position: Heidegger's Analysis of the Essence of Modern Technology," in *Heidegger on Technology*, ed. Aaron Wendland, Christos Hadjioannou, and Christopher Merwin (Routledge, 2018), 39-56.

14 Heidegger, "The Question Concerning Technology," 1.

(*poiei*) and *iv*) the final cause: what the artifact is used for, or its end.¹⁵

This definition conveys the essence of a technological artifact as the purpose it serves. Aristotle's final cause (*i*) suggests that the essence of a *techné* is simply its function. Indeed, this is an intuitive way to think about the purpose of a technological artifact, like a butter knife or a windmill, for example.

Heidegger argues that portraying technology's essence as a "mere means to an end"¹⁶ neglects its profound complexity. In particular, he believes Aristotle has overlooked the importance of *iii*): the maker cause.¹⁷ The Greek word for the production of a craft is *poiesis*. This word, however, was not just used in reference to technology but also for the creation of art and poetry. In Heidegger's view, the translation of *poiesis* as "production" is insufficient, and it is better defined as *bringing-forth*. According to Heidegger, *bringing-forth* is the *existential* enablement of technology.¹⁸

Bringing-forth is a deeply human process. It is a structure of existence that not only influences but actively reflects *being-in-the-world*. According to Heidegger, in the same way writing a poem reflects your worldview, the way we *bring-forth* technologies makes visible a quality of our being. This may seem like a jump. "Where have we strayed to?" Heidegger asks his reader, "What does the essence of technology have to do with revealing? The answer: *everything*."¹⁹ Heidegger traces the importance of "revealing" to the Latin term for truth, *aletheia*,²⁰ which more literally translates to unconcealment. The essence of technology, therefore, is the way it unconceals an aspect of *Dasein's being-in-the-world*. This might seem like a trivial claim because, according to Heidegger, a butter knife also reveals something about *being-in-the-world*. What is quite radical about Heidegger's argument is that he's not merely suggesting that technology is a particularly helpful concept to analyze, but that technology is *the locus* of revealing:²¹ the arena where our way of being is revealed in its entirety.

Heidegger elucidates this rather abstract argument with the example of the steamship. What is the essence of a steamship? A natural Aristotelian response would be that the steamship's essence is some combination of the materials it's made up of, the way it looks, and its purpose: traveling across the sea. Heidegger, on the other hand, views the essence of the steamship as what it reveals about *being-in-the-world*. Accordingly, Heidegger claims that the *existential* enablement of modern technologies like the steamship lies in what they reveal about our relationship to nature. Therefore, he suggests argues the essence of modern technologies like the steamship lies in what it reveals about our relationship to imperialism and globalization. In his words, the essence of the steamship is *enframing (gestell)*:²² the way it reveals how we view the natural world as a collection of resources at our disposal.

If the essence of technology is, as Heidegger says, what it reveals about *being-in-the-world*, what is the essence of LLMs? In other words, what do LLMs reveal about our way of being?

15 Heidegger, "The Question Concerning Technology," 1; Aristotle, *Metaphysics*, trans. W. D. Ross (Princeton University Press, 1984).

16 Heidegger, "The Question Concerning Technology," 1.

17 Heidegger, "The Question Concerning Technology," 4.

18 Heidegger, "The Question Concerning Technology," 5.

19 Heidegger, "The Question Concerning Technology," 5.

20 Heidegger, "The Question Concerning Technology," 5.

21 Dahlstrom, *The Heidegger Dictionary*.

22 Heidegger, "The Question Concerning Technology," 7.

IV. The Question Concerning Large-Language Models

LLMs are systems trained on vast amounts of data that provide responses to a wide variety of queries. According to Aristotle's final cause, the essence of a given *techné* lies in the purpose it serves. Accordingly, an Aristotelian definition might convey LLMs' essence as their capacity to democratize access to information. Just as the steamship orders oceans as navigable routes and continents as accessible markets, an LLM transforms vast repositories of text into accessible, contextualized knowledge tailored to individual needs. Heidegger would argue that in order to investigate the essence of large language models, it's necessary to begin by reflecting upon what they reveal about our *being-in-the-world*. In order to do this, it's first necessary to consider their *existential enablement*. What about the ontological structures of our existence makes the *bringing-forth* of large-language models possible?

When a user opens an LLM, they compose a prompt and receive a linguistic response. At first glance, this exchange appears merely instrumental: one may seek automation, summarization, drafting assistance, or task completion. Yet even these cases presuppose an orientation toward understanding. To summarize a text, to automate correspondence, or to generate code is not to bypass understanding altogether, but to redistribute or mediate it through language. It's evident, therefore, that the *existential* enablement of LLMs must concern our relationship to language and understanding. The leading question of this paper becomes: what do LLMs reveal about the way we use natural language to *understand*?

In *Being and Time*, Heidegger dedicates a term to this very concept—what he refers to as *discourse*.²³ Heidegger defines discourse as the process by which one's "understanding of being-in"²⁴ is disclosed to others through language. Importantly, Heidegger sees the *existential* enablement of the *ontic* existence of language as *discourse*. In other words, the particular letters on a page or the definitions in a dictionary are external manifestations of the *existential* structure of discourse. As Heidegger puts it, "the *existential ontological* foundation of language is discourse or talk."²⁵ *Discourse* doesn't literally refer to talk or conversations, but the very possibility of transferring *understanding* to one another. In technical terms, *discourse* is the *existential* possibility of successfully disclosing one's understanding. As a realm of revealing, LLMs must disclose something about our relationship to *discourse*.

V. The Flattening of Language: Heidegger on *Discourse* and *Idle Talk*

In order to make sense of Heidegger's perspective on *discourse*, it's first necessary to unpack what he means by understanding. What sort of understanding does effective discourse convey? Perhaps surprisingly, Heidegger argues that true (i.e., *authentic*) *understanding* is intimately grounded in phenomenological experience. *Authentic understanding*, he argues, can't be accessed completely through theoretical argumentation or logic.²⁶ This implies that understanding must have a pre-verbal, pre-theoretical dimension. In fact, Heidegger suggests that our most authentic form of understanding has a *pre-reflective* element: an aspect that precedes even conscious self-awareness.²⁷

The central component of our *pre-reflective* form of understanding, according to

23 Heidegger, *Being and Time*, 203.

24 Heidegger, *Being and Time*, 204.

25 Heidegger, *Being and Time*, 203.

26 Heidegger, *Being and Time*, 341.

27 Daniel O. Dahlstrom, *Heidegger's Concept of Truth* (Cambridge University Press, 2001), 192.

Heidegger, is our relationship to time.²⁸ The most authentic form of phenomenological understanding is the way we *project* our own *possibility* onto the world around us and the entities in it.²⁹ Heidegger’s term for this complex form of understanding is the *ready-to-hand*.³⁰ We understand something as *ready-to-hand* when we interpret it in terms of our own future. That is to say, we interpret it as equipment: what it can be used “as.” For example, let’s return to the example of the butter knife. Heidegger says we understand a butter knife as ready-to-hand when we actively interpret it as something that we can *use*. In other words, our most *authentic* way of *understanding* what a butter-knife is, is the way we pre-reflectively interpret it as something we can cut with. Heidegger sees the deepest form of human *understanding* as a relational activity, where the entities in front of us are endowed with meaning contextually in terms of time. Hereafter, I will refer to this phenomenological form of understanding as *primary understanding*.

Perhaps controversially, Heidegger argues that the ultimate purpose of *discourse* is to disclose *primary understanding* to others. However, a problem arises: to what extent is it possible to convey this rather abstract form of understanding to one another through words? Heidegger argues that *discourse* varies drastically in its effectiveness. In order to explain this, he makes an important distinction between *authentic* and *inauthentic* discourse.³¹ *Authentic discourse* is *discourse* that largely preserves one’s *primary understanding* of something. *Inauthentic discourse*, on the other hand, is *discourse* where one’s *primary understanding* is either drastically obscured or distorted.³²

Heidegger suggests that the fundamental cause of *inauthentic* discourse isn’t one’s general motivations or attitude—but a fundamental misunderstanding of the purpose, function, and limitations of natural language.³³ Controversially, Heidegger argues that all language is derived from our *primary understanding*. In his words, it’s impossible for a word or sentence to “disown its ontological origin,”³⁴ because all language is grounded in ready-to-hand understanding. However, words inevitably veil the complexity of the *primary understanding* they purport to express. This is because, according to Heidegger, we compress *primary understanding* into words by forming “assertions.”³⁵ One can imagine assertions as the way a gift (phenomenological understanding) is compressed so as to make it compact enough to wrap (express in words). According to Heidegger, when we construct *assertions*, our *pre-reflective* form of *understanding* is stripped of its temporal complexity as it “gets pushed into the uniform plane of that which is merely present-at-hand.”³⁶ The process by which understanding is stripped of its temporal complexity by turning into something *present-at-hand* (the opposite of the *ready-to-hand*)³⁷ is what Heidegger calls the “flattening” of language.

28 Heidegger, *Being and Time*, 203.

29 Heidegger, *Being and Time*, 183, 342.

30 Heidegger, *Being and Time*.

31 Heidegger, *Being and Time*, 341.

32 Heidegger, *Being and Time*, 341.

33 Although Heidegger’s doesn’t state this explicitly, this is clearly implied by his analysis of thematic discourse.

34 Heidegger, *Being and Time*.

35 Heidegger, *Being and Time*, 195.

36 Heidegger, *Being and Time*, 201.

37 *Present-at-hand* is what Heidegger refers to as the mode in which entities are encountered when they are considered theoretically or objectively, as isolated things with properties, detached from practical use and everyday involvement.

The concealing nature of assertions doesn’t suggest that it’s impossible to disclose primary understanding through discourse altogether, however. Although language inevitably conceals *ready-to-hand understanding* to a certain degree,³⁸ it’s possible for language to recover a significant portion of *primary understanding*.³⁹ According to Heidegger, the central difference between *authentic* and *inauthentic* discourse is whether language is used “thematically” or “unthematically.”⁴⁰ In highly *thematic* discourse, one takes the words in a sentence at face value. A thematic sentence is interpreted strictly according to the literal meaning of the words within it. Therefore, what is disclosed by the sentence is confined by the limitations of propositional truth. Somewhat controversially, Heidegger argues that the distortive process of forming assertions renders it *impossible* for *thematic discourse* to effectively disclose *primary understanding*. As Dan Dahlstrom explains in *Heidegger’s Concept of Truth*, “trying to recover the dynamics of this primary understanding by means of the logical structure of assertions or the parts of a sentence structure would be equivalent to trying to put Humpty Dumpty back together again.”⁴¹

In *unthematic* discourse, on the other hand, the subject of a sentence isn’t reducible to propositional truth.⁴² Although the distortive nature of assertions means *primary understanding* is invariably obscured to some extent, in *unthematic* discourse, other vehicles are relied upon to rediscover it. Heidegger argues that in *unthematic* discourse, *primary understanding* is primarily “made known in discourse and indicated in language by intonation, modulation, and the way of speaking.”⁴³ Our ready-to-hand understanding is expressed most effectively through the *way* something is said, not through mechanical propositional analysis.

Under this account, there is a clear hierarchy to how thematic—and thus, how *authentic*—various forms of discourse are. As Heidegger explicitly states, poetry preserves primary understanding to the highest possible extent. In poems, sentences are used extremely *unthematically*, as rhythm, tone, and metaphor disclose something altogether distinct from what is semantically encoded in them. On the other hand, an example of highly *thematic, inauthentic* discourse would be a Wikipedia article. In contrast to a poem, in online encyclopedia definitions, words are often used quite literally to convey meaning. Heidegger would likely argue that this is the case because Wikipedia articles are largely summaries of older online articles. Scientific articles cited by Wikipedia articles today have an average of 130.52 citations.⁴⁴ Crucially, whereas the words in a poem more directly aim to express an object of conscious experience (whether an emotion, or our *primary understanding* of a butter knife), most sentences in a Wikipedia article are constructed out of previous online sentences. Since each new sentence inevitably obscures *primary understanding* to a certain degree (whether they are thematic or unthematic), a sentence that takes another sentence as its inspiration inevitably distorts *primary understanding* more than a sentence that takes conveying *primary*

38 This is proven by Heidegger’s analysis of assertions in Section 33.

39 Heidegger, *Being and Time*, 206.

40 Dahlstrom, *Heidegger’s Concept of Truth*, 194.

41 Dahlstrom, *Heidegger’s Concept of Truth*, 192.

42 Dahlstrom, *Heidegger’s Concept of Truth*, 194.

43 Heidegger, *Being and Time*, 205.

44 Joshua M. Nicholson et al., “Measuring the Quality of Scientific References in Wikipedia: An Analysis of More than One Hundred Fifteen Million Citations to over Eight Hundred Thousand Scientific Articles,” *FEBS Journal* (author manuscript published July 1, 2020, final edited form November 19, 2020), PMID: 33089957, PMCID: PMC8060352.

understanding as its primary object. In consequence, as new sentences are repeatedly constructed out of older sentences to construct Wikipedia articles, *primary understanding* takes on less of the weight, and propositional truth takes on more. This engenders a pernicious cycle, where propositional truth constitutes a larger part of what is disclosed in a sentence, *primary understanding* is increasingly veiled, and as *primary understanding* is veiled, the more weight propositional truth carries.

Highly *thematic, inauthentic* discourse like Wikipedia articles is an example of what Heidegger would likely deem *idle talk*. In *idle talk*, discourse is interpreted completely “through what-is-said-in-the-talk, rather than through genuine primordial understanding.”⁴⁵ *Primary understanding* is entirely forsaken and forgotten, as sentences are only valued for their *thematic* meaning. As the pursuit of expressing *primary understanding* is entirely abandoned, *discourse* derails into a mere circulation of *thematic* word meanings. This leads to the rapid proliferation of sentences that no one phenomenologically “makes their own”⁴⁶ through *primary understanding*. Such ideas, in Heidegger’s terminology, are ultimately “groundless:”⁴⁷ devoid of true meaning.

The peril of *idle talk* is the way it encourages the continual reformulation and circulation of groundless ideas. As empty thematic sentences are used more often, they begin to “spread in wider circles and take on an authoritative character.”⁴⁸ The danger of *idle talk* is that sentences that are never phenomenologically understood are increasingly given credence. As such sentences proliferate, it becomes impossible to distinguish between what has in reality “been drawn from primordial sources and what is just gossip.”⁴⁹ Although Heidegger uses the word gossip here, his point likely pertains less to celebrity drama and more to philosophy students who use lots of complicated words they understand according to their semantic definitions. In *idle talk*, when we fail to acknowledge the limitations of natural language, we forget the value of *primary understanding* altogether.

The ultimate consequence of highly *thematic, inauthentic* discourse is way it bolsters the force of *Dasman* (i.e., *the they*.) *The they*, according to Heidegger, is an *existential force* that interferes with Dasein’s capacity to *understand* things *authentically*.⁵⁰ Many scholars interpret *the they* as referring to an anonymous, impersonal public or crowd that dictates norms, opinions, and behaviors in everyday life. The influence of *the they* further contributes to an obscurity of *authentic understanding*, because Dasein is tempted to yield their personal *primary understanding* in favor of what *the they* “says” or “thinks.” Crucially, *the they* is “nobody,”⁵¹ meaning Heidegger isn’t referring to a particular set of social norms, but more generally analyzing the process by which ideas gain social traction and authority. Heidegger conveys *primary understanding* as our ultimate mechanism of defense against the pernicious influence of *the they*. In consequence, as *primary understanding* is eroded through highly *thematic idle talk*, *the they* reinforces its domination. The radical consequence of *highly thematic discourse* is that, eventually and inevitably, “*the they* prescribes one’s state-of-mind,

45 Heidegger, *Being and Time*, 207.

46 Heidegger, *Being and Time*, 213.

47 Heidegger, *Being and Time*, 213.

48 Heidegger, *Being and Time*, 212.

49 Heidegger, *Being and Time*, 212.

50 Heidegger, *Being and Time*.

51 Heidegger, *Being and Time*, 166.

and determines what and how one sees.”⁵² Broadly put, Heidegger claims a highly *thematic* use of language serves to reinforce social norms. The radical aspect of Heidegger’s argument is his suggestion that the primary cause of *inauthentic* discourse, which fortifies *the they*, is an overreliance on propositional truth in our understanding of language (what he later refers to as “the logical prejudice”).

In Heidegger’s time, this theory was largely dismissed as overly vague and speculative. However, I believe, as a realm of revealing about our relationship to discourse, LLMs reinforce Heidegger’s view.

VI. A Realm of Revealing: Large-Language Models, *Idle Talk*, and *The They*

In this final section, I argue that as a realm of revealing LLMs, they reveal the veracity of Heidegger’s assessment of the “flattening” of language. I contend that: *i*) LLM’s formulation of words into vectors reflects Heidegger’s theory of assertions, *ii*) upon accepting *i*), LLMs training process resembles an incredibly *thematic* form of discourse, and *iii*) the qualitative mysteries of LLM’s high-dimensional vector spaces reflect the *existential* force of *the they*. Ultimately, I argue that LLMs instantiate *idle talk*. Therefore, as a realm of revealing, they reveal an already present mode of flattened discourse. In other words, I argue that the architecture of LLMs proves that the internal enablement of LLMs is a world already dominated by *idle talk*.

LLMs are trained on massive datasets collected from books, websites, articles, and other text sources to learn patterns in language.⁵³ These datasets often contain hundreds of billions or even trillions of words, drawn from diverse sources ranging from literary fiction to technical manuals, from historical documents to contemporary social media posts. By analyzing billions of examples of how often words appear together across these diverse texts and contexts, LLMs develop statistical representations of language patterns.⁵⁴ In other words, they statistically model patterns by tracking how frequently one word follows another. These linguistic patterns are quantified through their conversion into high-dimensional vectors—numerical representations in a space with thousands of dimensions.⁵⁵ For example, the word “flower” might be represented in vector space as “3492094290420,” where each number is a discrete dimension. This is easier to imagine in three-dimensional vector space, where each number in a vector represents a point on each of the three axes. At the beginning of the training process, each word’s vector denotation is assigned completely randomly.⁵⁶

The act of representing words with vector digits vividly reflects Heidegger’s theory of assertions. Recall Heidegger’s argument that we purport to express *primary understanding* in language by forming *assertions*: consolidating phenomenological complexity into something simple enough to express in words by conferring upon it a “definite character.” He claims that the temporal complexity of *ready-to-hand understanding* is inevitably veiled when we do this—the first step in the “flattening” of language. When he claims that assertions give *understanding* a “definite character,” Heidegger means that we strip *understanding* of tem-

52 Heidegger, *Being and Time*, 213.

53 Jay Alammar and Maarten Grootendorst, *Hands-On Large Language Models: Language Understanding and Generation*, 1st ed. (O’Reilly Media, 2024).

54 Alammar and Grootendorst, *Hands-On Large Language Models*.

55 Alammar and Grootendorst, *Hands-On Large Language Models*.

56 Alammar and Grootendorst, *Hands-On Large Language Models*.

poral aspect by reducing it to something present, concrete, and possible to externalize (i.e., write down). I contend that the representation of words as numbers vividly symbolizes this process. According to Heidegger’s definition, numbers seem to be as “definite” as something could possibly be. What could be more “definite,” “present-at-hand,” and feasible to write down than a numerical vector? Furthermore, the attribution of words with an arbitrary vector blatantly veils understanding (recall that although Heidegger sees words as distorting *understanding*, he doesn’t claim they are completely devoid of *primary understanding* altogether). In LLMs, whatever *primary understanding* is preserved in words is converted into a random number. It would be absurd to claim *any* sort of understanding, let alone Heidegger’s *primary understanding*, is intrinsic to these arbitrary number assignments. As in Heidegger’s theory of assertions, putting something into words strips understanding of its complexity; the transformation of words into vectors in LLMs obscures understanding completely. What does this aspect of LLMs reveal about our use of language? It seems the depiction of words as statistical digits reflects the way that assertions “flatten” understanding.

In *thematic* discourse, sentences are only interpreted according to the literal meanings of the assertions within them. The training process of LLMs physically embodies this process; words are only endowed with meaning according to the geographical representation of vector embeddings.⁵⁷ In the first stage of the process, a word is attributed with a completely arbitrary vector—which is analogous to the way assertions give *understanding* a definite character. LLMs are trained with continual mathematical reshaping and reforming of vectors. Each update involves calculating distances between vectors—measuring their dot products, their cosines, their geometric relationships—and shifting them incrementally to better predict co-occurrence patterns.⁵⁸ When a model incorrectly predicts the next word in a sequence, it adjusts the vectors slightly to improve future predictions. After processing billions of text examples through this optimization process, the vectors settle into configurations that maximize the model’s statistical accuracy.⁵⁹ Crucially, even more extremely than *assertions* in thematic discourse, vector assignments are completely empty of *primary understanding*. Thus, the only way LLMs form new sentences is by way of their mathematical representation in vector space (which is analogous to the meaning of the assertions—thematic meaning, in our analogy.) Assertions are the only basis for the production of new sentences, because the LLMs’ training process is devoid of true understanding. Therefore, their training process is an extraordinarily *thematic*, and thus radically *inauthentic* form of *discourse*.

Throughout this training process, each dimension (digit, or coordinate in vector space) is adjusted. Dimensions are sorted to represent a certain linguistic pattern, which can represent anything from “size” to “amplitude” to “offensiveness.” Each dimension is simply a numerical value—a coordinate in abstract mathematical space—with no inherent meaning outside the model’s internal system.⁶⁰ The model has no access to what “size” actually signifies; it only knows that certain tokens cluster together statistically. In fact, even experts cannot pinpoint the semantic meaning of each particular dimension. Although it’s possible to identify linguistic patterns, the difference between what dimension 2,393 and dimension 2,340 tracks remains mysterious.⁶¹ All that remains are linguistic patterns, but their precise

57 Alammar and Grootendorst, *Hands-On Large Language Models*.

58 Alammar and Grootendorst, *Hands-On Large Language Models*.

59 Alammar and Grootendorst, *Hands-On Large Language Models*.

60 Alammar and Grootendorst, *Hands-On Large Language Models*.

61 Alammar and Grootendorst, *Hands-On Large Language Models*.

meanings are unclear.

These dimensions are a vivid ontic representation of *the they*. Recall that *the they* is “nowhere:” it doesn’t refer to a particular social force, but the way ideas gain traction and credence through statistical social accumulation. Each dimension is given meaning by tracking linguistic patterns in books, websites, and articles, some of which are already deeply ingrained with linguistic patterns that reflect social norms. As these words are reduced to numerical digits, the socially influenced linguistic patterns are all that remain. As a meta-level mathematical representation of the way words tend to be used, LLMs are the ultimate embodiment of what “*they-think*” and what “*they-say*.” Literally, in LLMs, *understanding* is reduced to what word *they* tend to use after each word in a given context.

As a system that sorts and produces words by following linguistic patterns, LLMs are a potent representation of *idle talk*. By reducing understanding (words) to empty assertions (vectors) and continually passing them along through highly thematic discourse (repeated mathematical reconfigurations), LLMs completely forsake *primary understanding* in their production of novel sentences. This creates a system where sentences are perpetually reproduced and reformulated with no primordial *understanding* behind them. This leads to the constant “passing along” of *groundless* ideas. This perpetual recycling of empty word meanings ultimately leads to the reinforcement of what “*the they*” say: a bolstering of socially influenced linguistic patterns. According to Heidegger, this will lead to a world where *the they* (the linguistic patterns) dictate what and how we think.

The crucial conclusion of Heidegger’s “The Question Concerning Technology” is that LLMs don’t just symbolize, but actively reveal something about our way of being. The proliferation of LLMs isn’t just a bad example of discourse: it is a mirror. It reflects a world where our understanding of language has already been flattened. LLMs do not enable *idle talk*—*idle talk* is the *existential* enablement of LLMs. If, as Heidegger argued, a certain degree of human language hadn’t already been reduced to the passing along of *groundless* ideas we don’t understand, the success of LLMs would be impossible. In other words, their success proves that a significant part of our modern-day understanding of language has been reduced to *idle talk*.

Towards the end of “The Question Concerning Technology,” Heidegger suggests that technology not only reveals our way of being-in-the-world but actively shapes its course.⁶² Through *destining* and *enframing*, technology not only reflects but actively determines the way we relate to the world around us. Heidegger would likely predict that LLM models will constitutively reinforce our already flattened understanding of language. The *existential-ontological* implications of this continual process extend far beyond misinformation or the perpetuation of social biases:

“Ontologically, this means that as Dasein maintains itself in idle talk, it is—as being-in- the-world—cut off from its primordially genuine relationships-of-Being towards the world, towards Dasein-with, and towards its very Being-in. Such a Dasein keeps floating unattached, yet in doing so, it is always alongside the world, with Others, and towards itself.”⁶³

A world dominated by *idle talk* is a world in which we have forgotten to understand one another. In Heidegger’s language, as the bond between language and primary understanding erodes, authentic discourse will fall apart. Most of language will devolve into the constant rein-

62 Heidegger, “The Question Concerning Technology,” 9.

63 Heidegger, *Being and Time*, 214.

forcement of the ideas of *the they*: the meaningless regurgitation of groundless linguistic patterns.

VII. Conclusion

This paper has sought to question the essence of large-language models through Heidegger’s phenomenological framework. By investigating their *existential* enablement, I have argued that LLMs reveal something profound about our relationship to *discourse*. Their architecture—the conversion of words into vectors, the statistical reformulation of linguistic patterns, the perpetual circulation of assertions devoid of primary understanding—reflects the prevalence of highly *thematic, inauthentic discourse* in the modern world. As a realm of revealing, LLMs disclose the veracity of Heidegger’s claim that language has been “flattened.”

The proliferation of LLMs is not merely symptomatic but revelatory. It suggests that our understanding of language has already been reduced to *idle talk* to a considerable extent. If *primary understanding* had not already been largely forsaken in favor of propositional truth, the success of systems that operate purely through statistical patterns would be impossible. In this sense, LLMs are not the cause of idle talk but its manifestation—and perhaps its ultimate confirmation.

Whether LLMs will further entrench this flattening or prompt us to recover authentic discourse remains an open question. What is certain, however, is that questioning their essence has revealed something unsettling about our current way of *being-in-the-world*: a world where *understanding* is increasingly detached from phenomenological experience, and where *the they* maintains its pervasive influence over what and how we think.

Towards the end of “The Question Concerning Technology,” Heidegger offers us a glimpse of hope. Heidegger claims that the potential *saving power* from technology—which will enable us to reclaim a more authentic relationship with it—is a reclaiming of the act of *bringing-forth (poesis)*.⁶⁴ In Heidegger’s view, technological danger stems from the way it has corrupted the process of *bringing-forth*. Returning to the Greek term, *poesis*, which also translates to poetry, Heidegger suggests that a possible remedy for the corruption of language is poetic thinking. Clearly, LLMs can write poetry. However, on my reading, poetic thinking, doesn’t refer to poetry in the traditional sense—it designates a use of language that recaptures our primary, ready-to-hand form of understanding. Perhaps in this context, this paper, then, we can interpret him literally. Perhaps, the remedy to LLMs’ contamination of our understanding of language is a return to a more poetic way of speaking and thinking. In a world where LLMs reflect and encourage a flat understanding of language, perhaps we should turn to poetry to reclaim an *authentic* relationship with discourse.

Works Cited

- Alammar, Jay, and Maarten Grootendorst. *Hands-On Large Language Models: Language Understanding and Generation*. 1st ed. O’Reilly Media, 2024.
- Dahlstrom, Daniel O. *Heidegger’s Concept of Truth*. Cambridge University Press, 2001.
- Dahlstrom, Daniel O. “Im-position: Heidegger’s Analysis of the Essence of Modern Technology.” In *Heidegger on Technology*, edited by Aaron James Wendland, Christopher Merwin, and Christos Hadjioannou. Routledge, 2018.
- Dahlstrom, Daniel O. *The Heidegger Dictionary*. Bloomsbury Academic, 2013.
- Heidegger, Martin. *Being and Time*. Translated by John Macquarrie and Edward Robinson. Harper & Row, 1962.

64 Heidegger, “The Question Concerning Technology,” 14.

- Heidegger, Martin. “The Question Concerning Technology.” In *The Question Concerning Technology and Other Essays*, translated by William Lovitt. Harper & Row, 1977.
- Nicholson, Joshua M., et al. “Measuring the Quality of Scientific References in Wikipedia: An Analysis of More than One Hundred Fifteen Million Citations to over Eight Hundred Thousand Scientific Articles.” *FEBS Journal* (author manuscript published July 1, 2020; final edited form November 19, 2020). PMID: 33089957. PMCID: PMC8060352.

bell hooks: Abuse and Imperfect Love

Aidan Zhou

Abstract: What does it mean to love? I might respond one way, you might respond another, and bell hooks might respond that the asymmetry in our answers is exactly why we, as a society, have become so bad at loving. Love, despite its prevalence in both religious and secular conceptions of ideal societies, is surprisingly oft-neglected by many scholars in diagnosing the key issues afflicting contemporary society. Thus, in *All About Love: New Visions*, bell hooks urges us to consider love, or more so lovelessness (our inability to understand and practice love for one another), as what lies behind the widespread misunderstanding, conflict, and alienation we observe in modern society. hooks proposes a radical conception of love as defined by action rather than emotion, and as part of this redefinition of love, hooks suggests that love and mistreatment (what she labels abuse) can never coexist. In this essay, I argue that this particular claim about the coexistence of love and abuse does more harm than good for hooks' conception of love. I show that love does not *always* have to be perfect, and then I hope to show that at least some cases of abuse or mistreatment must be thought of as simply imperfect love. For love to preclude serious cases of abuse in the way she hopes for, hooks must rethink her definitions of love and abuse. Importantly, I do not hope to argue against love as a fundamental, relevant idea in ethics. Rather, by challenging hooks, and others, to critically re-examine the boundaries of loving action, I hope to resurface the idea of love as a practical ethic central to many of the most relevant contemporary issues that ethics is concerned with.

bell hooks begins *All About Love: New Visions* with a surprising premise: a majority of the world does not know how to love. Despite focusing significant cultural attention on the idea and practice of love, hooks argues that modern society suffers from the problem of lovelessness—we are raised in households absent of love, we are taught improper definitions of what love is, and as a result, we never learn how to properly love others.¹ To overcome this problem, hooks proposes that we adopt a 'love ethic,' a moral framework in which genuine love serves as an ethical standard for society. At the foundation of her love ethic, hooks presents a radically demanding conception of love, one in which "love and abuse cannot coexist."² In this essay, I reconstruct hooks' argument against the coexistence of love and abuse before showing that hooks must either accept that there are situations where love is

1 bell hooks, *All About Love: New Visions* (HarperCollins Publishers, 1999), 19.

2 hooks, *All About Love*, 6.

entirely impossible or make room in her definition of love for imperfect love: a form of love where one or more dimensions of love is absent. I then argue that many situations of abuse are structurally similar to imperfect love, and without a principled way of differentiating between them, hooks cannot claim that love and abuse can never coexist.

The primary feature of lovelessness, as hooks argues, is that we do not know what love means. We cannot begin to recognize, learn, and practice love in our everyday lives if we do not even have an understanding of what it *is*. Thus, borrowing from Scott Peck, hooks begins her argument by defining love: love is "the will to extend one's self for the purpose of nurturing one's own or another's spiritual growth."³ Here, there are two key features to note about hooks' definition. First, love for her is an *effective* will, one that disregards intention or feeling: "love is as love does."⁴ Love consists entirely of actions alone. Second, love must be aimed at some form of personal or shared growth; this growth can only be achieved if love contains certain constitutive dimensions: care, commitment, trust, knowledge, responsibility, and respect.⁵ hooks emphasizes that, for an action to be loving, it must contain *all* of these dimensions.⁶ After all, a major source of lovelessness is that love-like actions, actions containing some of these dimensions but lacking others, are often confused for love itself.⁷

It is worth noting briefly that this is a radical departure from traditional literature on love—Murdoch, Dover, Nehamas, and others all agree that love contains, on some level, an affective dimension; you feel love. hooks, however, chooses to define love as purely an action and independent of emotion, in part to encourage extreme responsibility over our behaviors.⁸ In such a way, we must view hooks' approach to love as part of a larger, ameliorative—not descriptive—project: hooks is concerned with love inasmuch as it informs our response to the problem of lovelessness. This will be important framing for later discussion.

hooks then goes on to argue that love and abuse can never co-exist; the presence of any form of abuse immediately precludes the possibility of love.⁹ Note that hooks is concerned with abuse because it can often appear love-like, that certain abusive actions can cause harm while still being delivered with care or genuine intent. Regardless, hooks believes these actions are still abusive and actively contribute to the problem of lovelessness—they trick us into thinking that harmful actions are loving. More specifically, hooks can be read as making one of two claims: the stronger claim that love and abuse cannot exist in a relationship or the weaker claim that love and abuse cannot exist in the same action. In this paper, I focus on the weaker claim about singular actions, for if I can show that love and abuse can coexist in a singular action, then I have also shown they can coexist in a relationship.

Before reconstructing the rest of hooks' argument, I will also clarify what exactly hooks is referring to when she talks about abuse. Abuse, for hooks, can be thought of as any *action* that involves domination, power, or control over another.¹⁰ This is an expansive, non-traditional conception of abuse. Nevertheless, hooks adopts this definition because it en-

3 hooks, *All About Love*, 4.

4 hooks, *All About Love*, 4.

5 hooks, *All About Love*, 7-8.

6 hooks, *All About Love*, 6-8.

7 hooks, *All About Love*, 5.

8 hooks, *All About Love*, 5.

9 hooks, *All About Love*, 6.

10 hooks, *All About Love*, 6, 41-42.

compasses certain non-paradigmatic forms of abuse like dishonesty, neglect, and betrayal that she is concerned with.¹¹ Lying, for example, is a form of abuse for hooks because it undermines the honesty necessary for loving relationships and results in harmful power dynamics.¹²

Importantly, however, hooks' *exact* positive conception of abuse is not essential to her argument against the coexistence of love and abuse. Rather, what matters is that, given a general conception of abuse, abuse is defined as never containing *all* of love's constitutive dimensions. Whether thought of as domination, power, or control, abuse will necessarily lack at least one of the core dimensions of love. Lying occurs in the absence of honesty, unkindness in the absence of affection, betrayal in the absence of commitment, and humiliation in the absence of care.¹³ We can intuit situations of abuse where only one dimension is missing—a compulsively dishonest parent whose actions are otherwise deeply love-like—or others where many dimensions are missing, as in the case of serious physical and verbal harm, but we cannot describe an abusive action that contains all of love's constitutive dimensions. Thus, hooks' argument hinges on defining abuse by this weaker, *negative* characteristic: in abuse, at least one of the core dimensions of love will be missing. For the rest of this paper, I will use "abuse" in this negative sense.

With this in mind, we can quickly synthesize hooks' argument that love and abuse are incompatible: love necessitates all of its core dimensions, and abuse lacks at least one of these dimensions. One action cannot be both love and abuse. While this might appear tautological, note that hooks is not arguing that abuse is *defined* by a lack of loving dimensions; rather, she only observes that abuse, as she defines it, always lacks at least one of these dimensions. In other words, hooks does not claim any non-loving action to be abuse, only that all abusive actions are non-loving. In premise-conclusion form, hooks' argument can be restated as such:

P1: Love is the will to extend one's self for the purpose of nurturing one's own or another's spiritual growth.

P2: This will to extend one's self for the purpose of nurturing one's own or another's spiritual growth manifests through actions that must contain certain constitutive dimensions: care, commitment, trust, knowledge, responsibility, and respect.

P3: If an action does not contain all of these dimensions, then it cannot be love.

P4: Abuse does not contain at least one of care, commitment, trust, knowledge, responsibility, or respect.

C: Abuse cannot be love.

In the next section of this paper, I object to hooks' formulation of love and argue that she has not adequately shown why love and abuse cannot coexist. I show that P3 is overly demanding and fails to account for general situations of *dimension-conflict*: when two or more dimensions of love are mutually exclusive. I then argue that if hooks accommodates love in general situations of dimension-conflict, which I believe she should, then she must also accommodate love in abuse.

To start, imagine that your friend Andrew asks if you enjoyed the gift he got you. You did not particularly enjoy it at all, but you also know for a fact that Andrew is a deeply

11 hooks, *All About Love*, 6, 42.

12 hooks, *All About Love*, 41.

13 hooks, *All About Love*, 6, 42.

sensitive person and that, no matter how caringly or gently you approach it, Andrew *will* be hurt if he finds out you did not enjoy the gift. On one hand, you might respond truthfully and inform Andrew that you did not enjoy the gift. By choosing to be truthful, however, you will be acting in violation of the knowledge that this will necessarily hurt him. On the other hand, you might act on your prior knowledge about Andrew's psyche and reply that you did, in fact, enjoy the gift. In this case, knowledge requires you to violate his trust through a dishonest response. While potentially trivial, and certainly non-abusive, I propose this scenario to illustrate the general possibility of situations where any possible action will lack a dimension of love. In the case of Andrew, you cannot act trustworthily without violating knowledge, and you cannot act with knowledge without violating trust. In situations like these, no possible action can contain all the dimensions of love, so it appears that no possible action can be loving.

To be consistent with her definition of love as requiring all of its dimensions, hooks would have to accept that, in situations like these, love is *impossible*; I argue, however, that hooks should not accept the impossibility of love. After all, love for hooks serves as the foundation of her love ethic, an ethical ideal she wishes to inform our actions. For love to function properly as an ethical principle, it should be informative, *especially* in these complicated situations of dimension-conflict. After all, an ethics that can only inform our actions in straightforward situations does not seem like a particularly salient kind of ethics at all. Additionally, while hooks could respond that love can still inform our actions inasmuch as we act in a way that best approximates love, this too feels counterproductive given our understanding of love as an ameliorative project: guiding us to perform more love-like actions would be unhelpful for hooks as she begins by claiming that love-like actions are a cause of lovelessness in the first place. Rather, if love is to serve as the ethical ideal that hooks intends, then she should want for love itself to always be possible.

hooks must then revise her definition of love to allow for love to be possible in all situations. In doing so, she would necessarily accept the possibility of *imperfect love*—love, which does not contain all of its constitutive dimensions. This is because if it is possible to be able to act lovingly in all situations, then it must be possible to act lovingly even in general situations of dimension-conflict. In these situations, love, whatever action it ends up being, will lack at least one of the key dimensions of love. Love in these cases will be imperfect, but it will still be love. Thus, P3 as stated cannot be true because love clearly does not *always* have to contain all of its dimensions.

In objecting to P3 and hooks' stringent requirement that love always contain all of its dimensions, I have also now objected to hooks' argument that love and abuse cannot coexist. For if imperfect love can still be love without all of its dimensions present, then hooks can no longer also claim that love and abuse cannot coexist solely on the basis of it lacking a dimension of love; love, as shown by imperfect love, does not always need all of its dimensions. In response, hooks must either find a way to differentiate between imperfect love and abuse, and explain why imperfect love is acceptable but abuse is not, or she must provide an entirely different argument as to why love and abuse cannot coexist—one that does not depend on the absence of love's dimensions.

In the final part of this paper, I will briefly consider a couple of ways in which hooks might attempt to differentiate between the absence of dimensions of love in abuse and the absence of dimensions of love in imperfect love. hooks might first claim that abuse is contextually different from the imperfect love allowed in general situations of dimension-conflict. After all, in situations of dimension-conflict like the case of Andrew and his gift, there exists

a tragic conflict between particular dimensions of love. A dimension of love must *necessarily* be absent in any resulting action. hooks could claim that abuse, however, is different from imperfect love because abuse doesn't contain a logical necessity for the absence of one of love's dimensions. Abusers make the *choice* to abuse when they could have acted perfectly lovingly otherwise. Thus, while abuse might carry the same structure as imperfect love, that of an action missing one of love's dimensions, the contexts in which these actions occur are drastically different. hooks might claim that what she really meant in P3 was that love must contain all of its constitutive dimensions, *whenever possible*. Such an amendment allows hooks to accept imperfect love only in situations where perfect love is impossible.

hooks is concerned with what can and often does arise from situations of tragic dimension-conflict where perfect love is impossible. Consider a father who is worried about his daughter's drug addiction. It has been a few months since she completed rehab, and he knows that any relapse would cause her severe physical and mental harm. In the past, attempts at reasoning or other supportive interventions have been unsuccessful. Knowing it is the *only option* to protect his daughter, he decides to closely monitor her every action: he reads her diary, checks her messages, and goes through her browser history. He tells her he does this out of love, but she finds his actions deeply controlling and abusive. At first glance, this appears to be abuse of the kind that hooks is concerned with—abuse that is complicated, somewhat love-like, and aimed at the well-being of the abused.¹⁴

This scenario involves, however, a structurally similar tragic conflict between two dimensions of love: the father cannot act in a way that honors his responsibility for his daughter without violating her trust, and he cannot honor his daughter's trust without sacrificing his responsibility for her well-being. He must choose between acting abusively and acting unlovingly. Note that the father is not *forced* to act abusively—he could, for example, choose to communicate honestly and openly with his daughter in a way that does not involve domination, power, or control. Rather, the tragic conflict arises because he does not have the option to act perfectly lovingly, to act in a way that contains all of love's dimensions. Because of this, hooks cannot use tragic conflict alone to differentiate abuse from imperfect love: both can come from contexts where perfect love is impossible.

Alternatively, hooks could claim that it is not context but rather the *extent* to which the dimensions of love are missing in abuse and imperfect love that differentiates them. I will not explore this possibility in great depth, but I note that for hooks to respond in this way would not be an easy task. The differences between imperfect love and abuse are often unclear in the real world, and it would be up to hooks to propose a way to principledly differentiate between the extents to which dimensions are lacking in imperfect love and abuse. In some examples, this difference might be obvious—an omission to a stranger versus physical violence against your child—but in others, especially containing abuse of the kind hooks cares deeply about, the differing extents to which the dimensions of love are absent might not be as straightforward. What she refers to as abuse, I might label imperfect love. If hooks really wants to argue that no form of abuse and love can coexist, and if she wants her argument to stay focused on abuse's lack of one or more dimensions of love, then the burden is on her to show exactly how such a lack of love's dimensions in abuse is principledly different from that of imperfect love.

Again, hooks could always choose to bite the bullet and accept that love is sometimes impossible. It might be easier to maintain that imperfect love cannot exist than to differentiate it from abuse. I have argued in this paper, however, that she should aspire to a more forceful

conception of love, one that can guide our actions even in complicated situations. In doing so, hooks would have to accept the existence of imperfect love and show us exactly how it differs from abuse; otherwise, hooks is wrong to proclaim that love and abuse cannot coexist.

Works Cited

hooks, bell. *All About Love: New Visions*. HarperCollins Publishers, 1999.

14 hooks, *All About Love*, 7.

Interpellation in the Colonial Context: Louis Althusser's Theory of Interpellation and Frantz Fanon's Account of the Race- Marked Colonial Person

Violeta Lopez Molina

Abstract: This paper examines the intersection of Louis Althusser's theory of ideology and Frantz Fanon's account of the race-marked colonial person to explore how interpellation functions—and fails—in a racialized world. Althusser's notion of interpellation describes how individuals are transformed into subjects under the ruling ideology through material practices. Yet while Althusser treats this process as universally reproduced under a given economic system and of “absolute guarantee,” Fanon reveals its limits: colonized subjects are not “always-already” interpellated as subjects. Drawing on Fanon's phenomenological descriptions, I argue that under white supremacist ideology, the black person is not constituted as a universal subject but as a race-marked person whose status as a human being is shaped by the colonizer's imposed imaginary. The default category of subject in the colonial order is the white subject, while the colonized person is treated as a “nonbeing.” Reading Fanon alongside Althusser exposes a structural asymmetry: ideology both constitutes subjecthood for some and denies it for others, producing what I term non-interpellation. Fanon's philosophy thus challenges Althusser's assumption of a shared field of subjectivation by showing that, in the colonial context, interpellation's “obviousness” relies on racial exclusion. Ultimately, I contend that confronting the colonial and racial limits of ideology requires rethinking the category of the subject—not as a fixed, reproducible universal, but as a historically contingent product of a material world divided between those who are interpellated as subjects and those who are denied subjecthood.

Introduction

In “Ideology and Ideological State Apparatuses,” Louis Althusser claims that “you and I are always already subjects, and as such constantly practice the rituals of ideological recognition, which guarantee for us that we are indeed concrete, individual, distinguishable and

(naturally) irreplaceable subjects.”¹ He takes a materialist approach to explain how ideology ensures the reproduction of the conditions of production—that is, how social institutions and practices secure the continuation of the existing economic system by ‘recruiting’ individuals into subjects that uphold it. However, he does not account for how the ruling ideology can reproduce itself not only through successful interpellation but also through its failures to interpellate certain individuals as subjects.² In this paper, I seek to bridge that gap in Althusser's theory of interpellation by bringing it into dialogue with Frantz Fanon's description of the black man's lived experience³ to claim that the race-marked colonial person is *not* constituted as a subject in the colonial context.

Situating this analysis primarily in the context of Western colonialism and its transnational afterlives across the African diaspora, I contend that the black person is not interpellated as a universal human subject like the (white) person but is instead viewed as a race-marked individual already overdetermined by their skin color. Interpellation under white supremacy is both reproduced and dialectically constituted by the ruling ideology of the colonizer, which denies black people interpellation as distinct, free subjects. Considering that the default category of the subject in the colonial context is the white subject, the colonized person is excluded from a universal form of interpellation. In this sense, Fanon's phenomenological account exposes a limit in Althusser's claim that “we are always-already subjects.”⁴ For those who are not properly interpellated as human subjects, interpellation fails not because the colonized exist outside ideology, but because ideology is structured to perpetuate their exclusion. This paper integrates the work of Althusser and Fanon—philosophers who developed interrelated concepts yet did not theorize them in direct relation—to show that interpellation for the black person is lived particularly (as race-marked), whereas for the white person it is lived universally (as unmarked). By introducing the concept of ‘non-interpellation,’ I claim that the black person can be denied subjecthood, revealing that the category of the subject is historically and racially delimited.

I. Althusser's Theory of Ideology and Interpellation

To clarify the framework I am using, it is important to distinguish Althusser's understanding of ideology from Karl Marx's. The main difference is that, for Althusser, there is no realm outside of ideology; ideology is not an illusory representation of real conditions⁵ or an abstract force external to reality, but a foundation of material existence. For Marx, ideology primarily functions as a set of false ideas imposed by the ruling class to obscure real economic conditions and relations of exploitation—what often gets called “false consciousness.”⁶ In

1 Louis Althusser, *Ideology and Ideological State Apparatuses*, trans. Andy Blunden (Monthly Review Press, 1971), 28.

2 I will be using ‘interpellation’ as a synonym for ‘subjectivation’ throughout the paper.

3 Fanon consistently refers to the “black man”. In this paper, I use “black,” “colonized,” or “race-marked person” interchangeably for clarity and inclusivity, while retaining his focus on race-marked colonial subjectivity.

4 Althusser, *Ideology and Ideological State Apparatuses*, 31.

5 By “real conditions,” Marx refers to the concrete material circumstances in which people live and produce through their activity under a given economic system—in this case, capitalism. He describes these as “real individuals, their activity and the material conditions under which they live” (Marx 1939, 3). Althusser uses the term in a similar sense to denote the material conditions structuring social life (i.e. employee-employer, citizen-state relations).

6 Engels first introduces the notion of “false consciousness” when he describes ideology as “a process

contrast, Althusser does not treat ideology as something people can think their way out of or examine from a neutral perspective. When he says that “men represent their real conditions of existence to themselves in an imaginary form,”⁷ he means that we live and act through images and narratives mediated by ideology; we are alienated because we do not recognize how our everyday material practices are shaped by, and in turn reproduce, ideology. Ideology is ‘imaginary’ because it provides a unifying story or set of images that makes a contradictory social reality intelligible and bearable in practice. In Lacanian terms, this corresponds to the imaginary order—the coherent image of self and world through which the subject is capable of navigating their surroundings.⁸ The imaginary belief that hard work leads to success, for example, shapes people’s expectations and behavior even though material realities (i.e., unequal starting points and structural barriers) contradict that idea. Ideology is thus imaginary not because it is fictional or false, but because it provides unified images and norms people internalize and enact, allowing daily life to perpetuate the status quo without confronting its underlying contradictions.

Ideology is also ‘material’ because its imaginary content only becomes effective when it is instantiated in commonplace practices. Attending school Monday to Friday, praying once a day, and clocking in at work are practices that transcend abstract beliefs. They reproduce the images and norms that make social relations intelligible, leading people to act on (ideological) representations because they cannot confront the “real conditions” directly. This makes alienation particularly effective: when subjects do not recognize how their everyday habits and self-relations are shaped by routineized social forms, they take those forms as natural, allowing ideology to reproduce the conditions of production more seamlessly. From a Lacanian angle, the imaginary provides a coherent, unifying set of images that makes such representations actionable. That coherence masks structural contradictions and lets people participate in and be subordinated by relations of domination without collapsing.⁹ The imaginary component of ideology would remain merely abstract and one-sided if it were not for its materiality. It is only through embodied, repeated practices that the ruling ideology is reproduced in everyday life.

Interpellation names the process by which ideology transforms mere individuals into sociocultural subjects. The relationship between subject and ideology is dialectical: ideology needs subjects to continue functioning, yet it produces those very subjects through interpellation, the process by which individuals recognize themselves—and are recognized—within ideology.¹⁰ Althusser argues that human existence fundamentally involves being a subject in-

accomplished by the so-called thinker consciously, to be sure, but with a false consciousness” (Engels 1893, 1).

7 Althusser, *Ideology and Ideological State Apparatuses*, 23.

8 Althusser’s account of ideology is psychoanalytically informed. Like Lacan, he sees the imaginary as shaping how subjects experience themselves and their world, giving coherence to social life while masking the structural conditions that produce it. Fanon also engages the concept of the imaginary and extends Lacanian theory to the colonial context (Fanon 1986, 161).

9 An example helps clarify this claim about how the imaginary prevents subjects from collapsing. Someone who has a stable sense of “who they are” in the eyes of others—say, a worker who can imagine themselves as a competent employee—can show up to their job, follow routines, and navigate authority without falling apart. Even if the workplace exploits them, the basic imaginary coherence (“this is who I am; this is how others see me”) keeps their world from disintegrating. Without that unifying self-image, daily interactions—being given instructions, responding to a supervisor, or understanding one’s place in a shared space—can become overwhelming or unintelligible, producing a kind of practical collapse.

10 Althusser, *Ideology and Ideological State Apparatuses*, 28.

terpellated under the ruling ideology since “ideology interpellates individuals as subjects.”¹¹ Interpellation, or hailing, happens whenever an individual is addressed by another. The simple act of hailing (i.e. “Hey, you there”) is an ideological operation which distinguishes just from inanimate objects and turns undifferentiated individuals into subjects part of the social order.¹²

Althusser distinguishes between two types of institutions that recruit subjects into ideology in society: the Repressive State Apparatus (police, army, courts) and the more discrete Ideological State Apparatuses (school, family, church, media, political parties, cultural institutions). Unlike RSAs, which function primarily “by violence,” the ISAs operate “massively and predominantly by ideology.”¹³ Althusser points to how the school in particular emerges as the dominant ISA under capitalism, replacing the church in its role of reproducing the relations of production. Children, he notes, are “obliged to attend the school, which takes children from every class at infant-school age and then keeps them there for years.”¹⁴ In the process, they are not only taught skills but also the rules of good behavior and the values necessary to reproduce the social order. This material dimension of ideology becomes evident in everyday rituals and institutions: a mother calling her child “son” within the family ISA, a teacher marking attendance in the school ISA, attending a birthday or religious ceremony through the cultural or religious ISAs, queuing at a store within the economic ISA, or greeting someone with a smile or handshake as part of the civic ISA. Althusser views these as ideological acts through which subjects are formed and the capitalist order can be reproduced. The plurality of ISAs attenuates and extends subject recruitment without relying solely on state violence. Institutions are built on bourgeois norms and interests, which are propagated through either ideology (ISAs) or, if necessary, coercion (RSAs). If, for instance, an individual refuses to turn around when hailed by a police officer, only then does the RSA intervene to ensure compliance. In this sense, Althusser extends Marx’s view that capitalist ideology reproduces itself not only through the state’s economic base but across the self-evident spheres of everyday life.

Because it is not a temporal, two-step process, interpellation is not sequential; the existence of ideology and hailing are almost like two sides of the same coin. Althusser’s claim that “individuals are always-already subjects”¹⁵ suggests that we are born into an ideological structure that interpellates us before we can recognize ourselves. Interpellation occurs through material, everyday, ideological rituals present even before birth,¹⁶ manifesting in the name, gender, household, and school one is assigned. Its dialectical character also makes it efficient, for ideology would not function if subjects did not treat social forms, names, or expectations as ‘natural.’ Although interpellation may seem obvious, it is an operation simultaneously reproduced and enabled by ruling ideology and its specific category of the subject.

II. The Universal and the Embodied Particular in the Colonizer’s Imaginary

But what happens when an individual is treated not as a subject, but as an object? Fanon shows how the ‘obviousness’ of interpellation diminishes for the colonized subject

11 Althusser, *Ideology and Ideological State Apparatuses*, 28.

12 Althusser, *Ideology and Ideological State Apparatuses*, 30.

13 Althusser, *Ideology and Ideological State Apparatuses*, 16.

14 Althusser, *Ideology and Ideological State Apparatuses*, 18.

15 Althusser, *Ideology and Ideological State Apparatuses*, 30.

16 Althusser, *Ideology and Ideological State Apparatuses*, 31.

living under the colonizer’s ruling ideology. He criticizes the presupposition of a symmetrical subject-subject relation by illustrating how the black person is not interpellated as a subject in the colonial context. The example where Fanon hears a kid telling his mom, “Look, a Negro!”¹⁷ on the train illustrates a radically different experience from Althusser’s direct, general address, “Hey, you there.” The kid does not address the black man directly, but indirectly, through his mom. The addressive subject-subject relation occurs only between two colonists; the race-marked individual is constitutively excluded from reciprocal face-to-face interpellation despite clear presence. Fanon finds himself treated like a toy on a shelf or a museum exhibit—as an “object in the midst of other objects.”¹⁸ The fact that the most basic act of hailing that Althusser describes is not lived as a universal human experience for the black person in this scenario reinforces the dialectical relationship between ideology and interpellation in a white supremacist world. Instead of being interpellated as a subject, the colonized person is treated as a race-marked quasi-object, having interpellation refused because it is structured by the imposed story of the colonizer’s ideology.

This “story” goes back to what both Althusser and Fanon understand through the Lacanian imaginary. Since the imaginary refers to how people form a sense of self by identifying with external images, in a colonial context, this means the colonized are forced to live within the colonizer’s imagination. They must comply with how the world is made intelligible to the colonizer. The ruling ideology—the settlers’ ideology—identifies and assimilates white individuals as constituting the category of the subject while perceiving the colonized, race-marked person as “the unidentifiable, the unassimilable”¹⁹ *other*. The colonized see themselves through distorted images created by the colonizer, internalizing self-images imposed on them. For the black person, interpellation and ideology are not lived spontaneously or subtly as they are for the white/colonizer. Instead, the historical and economic practices of a foreign civilization structure every encounter with its norms and images, producing a form of alienation experienced both psychically and bodily.²⁰ From this perspective, Fanon expands Althusser’s account of the imaginary by showing how, in the colonial context, it is a site of violent distortion, where the black person experiences their body and actions through the colonizer’s imposed story.

Fanon notes, “In the white world, the man of color encounters difficulties in the development of his bodily schema. Consciousness of the body is solely a negating activity.”²¹ Here, gestures, movements, and glances are read through the colonizer’s imaginary gaze, making the black person experience their body as an unfree object. For instance, a white pedestrian walking alone at night might move “naturally,” their movement expressing the taken-for-granted alignment between their body and the world that ideology interpellates as “obvious.” A black pedestrian, by contrast, is more likely to anticipate the racialized gaze, adjusting posture, pace, and gestures, asking, “Will this confirm the stereotype?” Even simple acts of moving through space are mediated by ideology, with the body not acting spontaneously but adjusting in response to the white imaginary, which seals the black person into “crushing objecthood.”²² Rather than being interpellated as full subjects, colonized people

17 Frantz Fanon, *Black Skin, White Masks*, trans. Charles Lam Markmann (Pluto Press, 1986), 109.

18 Fanon, *Black Skin, White Masks*, 109.

19 Fanon, *Black Skin, White Masks*, 161.

20 Fanon, *Black Skin, White Masks*, 13.

21 Fanon, *Black Skin, White Masks*, 110.

22 Fanon, *Black Skin, White Masks*, 109.

experience movements as objectified, measured against the dominant gaze. While Althusser shows how subjects are hailed by everyday material practices as subtle domination, Fanon complicates this: when interpellation is refused, it becomes disruptive rather than discreet. Colonized individuals live subjecthood as a “negating activity,” continually negotiating between externally imposed images and embodied experience. In this way, colonial ideology operates both in the realm of the imaginary—where recognition, both of oneself and by others, depends on the colonizer’s imaginary—and materially, through institutions that shape bodies, routines, and social relations in ways that reproduce that same imaginary through interpellation.

In the colonial context, both “Hey, you there” and “Look, a Negro” respond to acts guided by the same ruling ideology: white is the default, abstract, and universal category of the subject. The universal, presupposed category of the subject is fundamentally white, whereas black-skinned persons are particularized and epidermically “overdetermined from without.”²³ Paradoxically, white supremacy is an ‘abstract’ type of universality because it is not objectively universal; it is particular in the material world, since not everyone is white, yet it perceives itself as universal and interpellates others in its terms. If it were objectively universal, it would not heterogenize subjects by skin color. Human subjects act under the misconception that ideology is illusory, despite their actions being based on ideological terms and racial biases, such as when a white person is alarmed by a black pedestrian. Under this ruling ideology, whiteness is internalized, reproduced, and treated as rational and civilized in contrast to the black person, who is conceived and treated as animalistic, violent, inadequate, and “made of the irrational.”²⁴

Althusser’s universal structural description of what it is to be human is experienced as particular by the colonized person, who is not interpellated as a (universal) subject but as race-marked, defined always in relation to the white person: “dissected under white eyes, the only real eyes.”²⁵ The sociocultural unconscious assumes the superiority of white over black. Although Althusser’s example of interpellation (“Hey, you there!”) might seem non-ideological on first impression, Fanon’s train experience—where interpellation is denied—highlights the material dimension of ideology, instantiating the dominance of the white imaginary.

III. The Limits of Althusser’s ‘Always-Already’ Subject

The constitutive character of ruling ideology explains why black persons cannot easily assert themselves as universal, free subjects in the colonial context. Fanon’s experience aligns with Althusser’s claim that what seems to happen *outside* ideology—on the train, for example—actually occurs *within* it.²⁶ The colonized subject cannot escape ideology; its exclusion from subjecthood is reinforced by material circumstances in which the supremacy of white values “imposes a dichotomy upon the whole people.”²⁷ Ideology, like the unconscious, shapes everyday practices so thoroughly that it becomes invisible. Fanon points to how even after the constitutional abolition of slavery in France, the colonial race-marked person could not assert himself as a free self-consciousness.²⁸ The black person was freed by his master

23 Fanon, *Black Skin, White Masks*, 116.

24 Fanon, *Black Skin, White Masks*, 123.

25 Fanon, *Black Skin, White Masks*, 116.

26 Althusser, *Ideology and Ideological State Apparatuses*, 31.

27 Fanon, *Black Skin, White Masks*, 45.

28 This resonates with Du Bois’s concept of double-consciousness: the Black individual sees himself not

and did not fight for this freedom, leaving interpellation bound to white-defined norms.²⁹ As a result, justice and liberty remain white justice and white liberty rather than genuinely universal values.³⁰

Following Althusser, the Repressive State Apparatus does not necessarily transform the Ideological State Apparatuses that continue to reproduce racial subjection. Schools, the family, religion, and culture sustain the white image of the subject even when formal domination appears absent. For interpellation to operate as an undifferentiated process of subjectivization, an entire material transformation of the category of the subject would be required. If Althusser insists that emancipation from ideology is impossible, then a genuinely humane interpellation demands a radically new world—one that institutes an objectively universal, non-racist category of the subject. Even rebellion does not free the race-marked subject so long as the institutions reproducing ideology (ISAs and RSAs) continue to hail individuals in the settlers' terms. Fanon emphasizes that being black does not guarantee awareness of this structural problem: aspiring to whiteness or reacting with hatred to it are complementary responses to how the colonized are denied adequate interpellation.³¹ Colonial ideology thus traps the colonized between trying to fit the colonizer's imaginary and frustration at being deprived of interpellation, keeping them bound to the settler's framework.

The implication is that in the colonial context, interpellation is twofold: the world is divided between colonizers and colonized, interpellated and non-interpellated. Althusser's "Hey, you there" represents the ideal condition of interpellation, yet Fanon provides an example of how interpellation concretizes under colonialism while highlighting its dehumanizing dimension. Colonization imposes a forced recategorization of the subject, constituting the colonized only in relation to the settler. Althusser treats interpellation as an act of domination, whereas Fanon emphasizes that some are denied interpellation altogether, which also manifests relations of domination. Althusser presumes that hailing succeeds in constituting subjects and primarily addresses class reproduction, but he overlooks how race and colonialism can obstruct subject-formation, and how such failures themselves reinforce ideology.

Fanon's phenomenology offers insight into what Althusser's framework does not account for: the denial of interpellation for race-marked subjects. For Althusser, ideology operates through recognition and misrecognition within the imaginary. Subjects recognize themselves as interpellated, misrecognizing this recognition as self-standing agency.³² Misrecognition is not being mistaken for the wrong person but mistaking one's ideological positioning for genuine autonomy. In the familiar scene of interpellation, a passerby responds to "Hey, you there!" and is inserted into the imaginary relation ideology requires. Fanon's train

fully through his own self-consciousness, but constantly through the eyes of the dominant group, measuring his worth and existence against a world that denies him full interpellation. As Du Bois writes, the black person is "born with a veil, and gifted with second-sight... One ever feels his twoness, —an American, a Negro; two souls, two thoughts, two unreconciled strivings" (W.E.B. Du Bois 1903, 3). In the colonial context Fanon describes, a similar dynamic operates: the Black subject is denied interpellation as a full subject, forced to negotiate identity and bodily existence within the imposed values of the white master.

29 Fanon, *Black Skin, White Masks*, 219. Fanon explicitly draws from Hegel's dialectic of recognition, citing that "man is human only to the extent to which he tries to impose his existence on another man in order to be recognized by him" (216). Yet unlike Hegel, Fanon exposes how colonial relations foreclose mutual recognition, making the very conditions of reciprocity—the foundation of human freedom—historically impossible.

30 Fanon, *Black Skin, White Masks*, 217.

31 Fanon, *Black Skin, White Masks*, 10.

32 Althusser, *Ideology and Ideological State Apparatuses*, 29.

example, by contrast, illustrates being excluded from interpellatory processes since he is not hailed as a self-conscious subject but objectified, "abraded into nonbeing."³³ This appears as an illustration of non-interpellation that cannot be explained through Althusser's framework. The child does not expect Fanon to "turn round" because he does not view him as a subject in the first place. The colonized are necessary to the system yet subordinated; their labor and bodies are exploited to sustain the world that denies their subjecthood.

It is important to note that Fanon does not aim to recover a pre-existing essence of black selves but insists on an active struggle for what one might, in Althusserian terms, call humane interpellation wherein the colonized must "make himself recognized."³⁴ Fanon's expression "All I wanted was to be a man among other men"³⁵ captures the structural impossibility of adequate interpellation within a racialized world. His experience³⁶ shows the colonized are barred from subjecthood and that this exclusion turns subject formation into a struggle to exist otherwise. Interpellation is thus a fight against the colonizer's ideology: "I demand that notice be taken of my negating activity insofar as I...do battle for the creation of a human world—that is, of a world of reciprocal recognitions."³⁷ Fanon reframes subjectivity as a contested field in which interpellation must be wrested from historical structures that deny it by making it seem natural, showing that Althusser's "immutable" subject³⁸ is neither neutral nor universally experienced.

Putting Althusser and Fanon into dialogue raises the question: what form of being is available for those excluded from subjecthood? While Althusser helps explain how ideology constitutes subjects through interpellation, Fanon implies that it also reproduces itself through *non-interpellation*—through the systematic refusal to constitute certain individuals as subjects. In this paper, I have argued that non-interpellation is not a failure of ideology but part of its function. It sustains the ruling order by determining who is recognizable as human in the first place. Fanon's account of racialized exclusion thus exposes the limit of Althusser's claim that "we are always-already subjects," further suggesting that interpellation's universality is fractured by the colonial and racial conditions that underwrite it. Fanon's vision of "a world of reciprocal recognitions"³⁹ opens onto the possibility of reconfiguring ideology itself, toward a form of subjectivity grounded not in the reproduction of whiteness as universal, but in the interpellation of all humans as subjects with distinct histories and equal humanity.

Works Cited

Althusser, Louis. *Ideology and Ideological State Apparatuses*. Translated by Andy Blunden. Monthly Review Press, 1971.

Fanon, Frantz. *Black Skin, White Masks*. Translated by Charles Lam Markmann. Pluto Press, 1986.

Du Bois, W. E. B. "The Forethought." In *The Souls of Black Folk*, 1903.

33 Fanon, *Black Skin, White Masks*, 109.

34 Fanon, *Black Skin, White Masks*, 217. Fanon does not use the term 'interpellation' himself; I am arguing that his account of the struggle for recognition can be read as a fight to be properly interpellated.

35 Fanon, *Black Skin, White Masks*, 112.

36 The suggestion that his personal experience is not merely personal is evident when he writes "I was responsible at the same time for my body, for my race, for my ancestors" (Fanon 1986, 109).

37 Fanon, *Black Skin, White Masks*, 218.

38 Althusser, *Ideology and Ideological State Apparatuses*, 22.

39 Fanon, *Black Skin, White Masks*, 218.

Engels, Friedrich. "Engels to Franz Mehring." In *Marx-Engels Correspondence 1893*. New International, 1934. https://www.marxists.org/archive/marx/works/1893/letters/93_07_14a.htm.
Marx, Karl, and Friedrich Engels. *The German Ideology*. International Publishers, 1939.

Psychological Influence on Deceptionist Lies

Francesca Stamati

Abstract: This paper examines the psychological dimensions of lying within the deceptionist framework, proposing that traditional theories that focus solely on intent to deceive fail to capture the complexity of why people tell lies. I primarily examine Jennifer Lackey's revised deceptionist account of lying, which broadens deception to include both explicit falsities and acts of concealment. Lackey's distinction between deceit and deception appreciates some of the nuances of lying, expanding other, more modest deceptionist theories. However, Lackey's view still overlooks lies that cannot solely be explained by conscious intent. Thus, to address these flaws, I propose a refined deceptionist view that considers both conscious and subconscious influences. Both forces may be responsible for everyday "dishonest" speech acts that are driven by the desire to evade internal conflict or emotional discomfort, not to mislead others. By integrating insights from psychology—especially the roles of self-deception and cognitive dissonance—this paper reimagines lying as a complex speech act that mediates between conscious assertion and subconscious self-preservation. Ultimately, I argue that a deceptionist theory informed by psychology more closely accounts for the blurred boundaries between lying to others and lying to oneself.

Traditionally, language philosophers have prioritized deceptive intent to distinguish lying from other types of non-alethic discourse, such as bullshit or falsity. Whereas traditional deceptionists consider intent to deceive to be a necessary condition of lying, non-deceptionists do not and have provided several accounts against this view. However, in this paper, I will consider only the deceptionist perspective. In particular, I will examine and develop Jennifer Lackey's revised deceptionist approach, which extends the traditional view to acknowledge the disconnect between personal belief and assertions. However, this view still fails to account for shame-motivated lies, self-deception, and subconscious intentions. Therefore, a deceptionist theory that considers the many practical functions of lies as speech acts would better appreciate the profound nuance of lying from a psychological perspective.

By elaborating on the definition of deception, Lackey addresses more cases of lying than pre-existing views. The traditional deceptionist approach applies the following princi-

ples to categorize a lie. Under this view, *A* lies to *B* about *p* iff:

- I) *A* states that *p*
- &
- II) *A* believes that *p* is false
- &
- III) *A* intends to deceive *B* in stating that *p*.

In “Lies and Deception: An Unhappy Divorce,” Lackey notes that this traditional theory employs a limited definition of deception, leaving it vulnerable to non-deceptionist rebuttals. However, she still maintains a version of deceptive intent in her revision. To broaden the scope, Lackey distinguishes ‘deceit’ from ‘deception.’ While deceit is marked by an intent to inspire a false belief, Lackey says, deception involves an intent to conceal information or evidence.¹ In her adaptation of deceptionism, Lackey incorporates her broader view of what it *means* to be deceptive in principle III:

- I) *A* states that *p*
- &
- II) *A* believes that *p* is false
- &
- III) *A* intends to be deceptive to *B* in stating that *p*, meaning that *A*
 - i) intends to deceive
 - or
 - ii) intends to conceal information or evidence.

This revision regards both explicit falsities and the act of keeping secrets as lies. In her third condition, Lackey alludes to Paul Grice’s work on conversational norms, which outlines four maxims for cooperation: Quantity (say just enough), Quality (say what you believe to be true), Relevance (be relevant), and Manner (be clear, brief, orderly, etc.). While traditional deceptionists describe lies as violations of Quality in principle III, Lackey annexes the perversion of Relevance (ii) via concealing pertinent content as another form of deception. This distinction is significant because it includes more covert forms of lying, such as cases where the speaker intentionally avoids sharing relevant information, though they may never explicitly contradict that information with what they *do* say. As the act of keeping secrets is often considered a lie in colloquial use, Lackey’s expansion feels more natural and aligned with what we intuitively perceive as lies. Suppose a child steals a cookie from the kitchen, knowing her father would reprimand her if he found out. When her unknowing father casually asks his daughter what she did that afternoon, she avoids any talk of cookies. While the traditional approach would not consider this a lie, the child’s father might argue otherwise. This concealing of information, while covert, is intentional and active—and still a lie. Lackey’s version rightfully includes such a case.

Furthermore, Lackey defends deceptionism by proposing the concept of selfless assertion—the separation of a speaker’s utterance from their internal beliefs in certain contexts. If the intent to deceive is necessary, then these assertions do not count as lies, an exclusion she regards as appropriate. According to Lackey, when one’s beliefs contradict their spoken statements, this does not indicate that the speaker is lying. Suppose the cookie-thief’s father is devoutly religious, but he chooses not to inflict his theological views on his daughter. He

¹ Jennifer Lackey, “Lies and Deception: An Unhappy Divorce,” *Analysis* 73, no. 2 (April 2013): 241, <https://doi.org/10.1093/analys/ant006>.

may believe, wholeheartedly, that stealing is a sin that immediately warrants punishment in Hell. However, he does not tell his daughter this, instead putting her in timeout and telling her “It’s OK to make mistakes”—suggestions he read in a book on parenting skills. Despite his religious beliefs, he chooses not to scare his child with damnation, instead reassuring her with a parenting cliché he may not even believe. Is his utterance a lie? Lackey would argue, no; he does not aim to inculcate a false belief in his daughter. It may be that the father trusts the authority of his parenting books and the statistical results of studies on reprimanding children in a healthy way. If the father believes he *is* lying, then he might also believe he is committing a deadly sin that will land him in Hell. This religious absolutism would be rather problematic and restrict him to act only under one viewpoint in all contexts to retain his piety. Instead, he considers his personal religious beliefs irrelevant to the context of parenting his child. Selfless assertion reaffirms the need for an intention to deceive; otherwise, such cases that lack active intent are called lies.

Lackey’s proposal permits a more liberal and intuitive coexistence of personal belief and spoken words. There are many situations in which we *do* separate our personal biases from fact, acting as a vessel for the views of a third party—even an imaginary one. A ripe example is debate club. Debaters must be able to advocate for both sides of any argument, regardless of their own position on the matter. Consider a pro-life female debater arguing for abortion. While she may not agree with the perspective, she can still effectively support her claim using hard facts, reason, and other rhetorical devices. We might question her allegiance to the pro-life cause, but we would not declare that she is *lying* by participating. Instead, Lackey would say, the debater simply temporarily suspends her own beliefs, assuming the role of someone with another informed perspective. This act of simulation is allowed under Lackey’s principle of deceptive intent, loosening the connection between lying and negating personal belief.

These interpretations of particular scenarios of lying are compatible with Lackey’s theory. Commonly, people *do* lie by concealing information, supporting her view. Some situations, however, reveal that this perspective fails to cover all cases of lying. Let us remember the child who stole a cookie from her father, intentionally omitting this act later when telling him about her day. Assume this child concealed her actions solely because she is embarrassed to admit to them (like many children do). Even if the father knows she had taken a cookie, the daughter might still explicitly lie—screaming, “I didn’t do it!” Her utterance serves not to convince the father, but to avoid the shame of admission. Lying plainly out of embarrassment does not seem to intend to conceal evidence, nor does it serve to motivate a false belief in the hearer. It is a speech act intended to protect one’s dignity by denying an act they consider morally ‘wrong.’ One must take a roundabout, unnatural path to understand such a lie in terms of concealing evidence. It would seem, then, that Lackey’s definition of being deceptive must expand to include other motives for lying, such as to avoid embarrassment or other negative feelings resulting from confession.

An apt, yet complex example of another motive is self-deception—a psychological phenomenon that Lackey and other philosophers erroneously neglect. Certain utterances can be a vehicle for lying to the self, using an interlocutor to aid in the process. At the psychological level, it appears that the speaker is motivated to influence not the hearer’s cognitive state or beliefs, but their own. While intention is unclear, it still interacts subconsciously. This nuance is visible in cases that appear to fit Lackey’s idea of selfless assertions, yet still covertly involve conflicting beliefs. Consider the father who ‘suspends’ religious fervor to be a good

parent. He appears to suppress his theological beliefs in favor of expert-recommended parenting strategies. But this swapping still involves ideology. The father is simply allowing his other belief—the belief that he should obey parenting books to be a good parent—to override his religious beliefs. Neither belief detaches or disappears from the speaker’s utterance; one is simply smothered via the act of lying. Why *should* the father trust parenting books? Plenty of parents do not. The difference is in their individual beliefs in the validity of these third parties. Even if this third-party source is widely trusted and reputable, the father’s commitment to it still involves belief.

Therefore, Lackey errs in eliminating the self from speech. The father’s act is not truly selfless; if the father were to claim that it *is*, this might be a self-deceptive tool to protect himself from his clashing, irreconcilable values. In this regard, the lie acts as a mechanism that smooths over the rough internal disjunction between personal beliefs. It may not be possible for a human—informed by other humans through the lens of ideology and opinion—to make a truly ‘selfless’ assertion. Moral values and cultural norms frequently conflict, even without us knowing. It would appear, then, that the father *is* lying in some sense by favoring one belief over another; he may be lying to a *part* of himself. Although it is mostly unintentional and unconscious, this speech act is still a variation of deceiving the self.

Self-deception happens all the time. Our conscious beliefs and values often contradict our actions and utterances, and the real lie is that both polarities can coexist. So, what is the purpose of lying to the self? Often, it allows speakers to act immorally while maintaining a clear conscience or to make an exception of the self that contradicts principles they typically apply to others. However, it can also serve to resolve other internal tensions. While this mechanism fires up especially for core beliefs involving religion and culture, it is frequently, if not always, present to some degree. There are many examples that are more obvious than the father’s case. Self-deception is how people who value a healthy lifestyle still lounge in bed watching TV and eating Dominos. Speech acts are often used to resolve these contradictions, with assertions such as “One more pizza slice can’t hurt.” The speaker may consciously disagree with that statement two hours later, when they are bedridden with a stomachache. Their utterance can then be considered a lie, which they may have consciously believed at the time, but subconsciously experienced as a contradiction to their belief that a healthy lifestyle means *not* gorging on pizza. This disconnect between belief and behavior—often called ‘cognitive dissonance’—is a common example of lying to oneself.

Ultimately, Lackey’s concept of lying fails to address such cases. However, self-deception is substantive to discourse regarding lies as a speech act. When a speaker utters a statement they temporarily *convince* themselves to be true, the person claims this imaginary ‘other’ perspective—that which was misused by Lackey’s selfless assertion—as their own, even if it clashes with other beliefs. For the father, this occurs the moment he assumes the point of view of those writing his parenting books. He does not suppress his religious beliefs but rather overrides them with a lie. When a speaker’s conscious beliefs converge with another perspective, yet that person subconsciously still retains opposing beliefs, the lie covers this gap in belief. It can then be said this speaker is deceiving themselves—and the utterance serves as a speech act to do so. This nuance is not appreciated by Lackey’s approach, which neglects the element of covert deception in both the second and third principle. She would consider cases like the father’s utterance to be truthful, a fault that must be resolved in our revised theory of deception.

So how do we accommodate to this kind of lie? To respect the omnipresence of belief

in our utterances, we must specify the definition of ‘belief’ in our conditions. As the term can have great nuance, with multiple beliefs colliding at different levels of the speaker’s awareness, this is a difficult point to capture. Additionally, as self-deception can be totally, partially, or subconsciously, one cannot neatly assign the verb ‘intend’ to such a case. When someone makes a statement contradicting their true beliefs in favor of another belief set, consciously believing the two align, they do not *intend* to deceive. By definition, they do not even consciously believe themselves to be lying. But we still may say that unaware speakers like the father are “lying to themselves.” Furthermore, ‘intention’ is quite obscure across a broad range of lies, as people often do not consciously know *why* they are lying. Compulsive liars are among these cases. Those with this behavioral issue habitually lie without any concrete reason; they may not truly *intend* to deceive anyone. Rather, these speakers make random false statements on impulse, possibly to avoid the discomfort of resisting this compulsion and telling the truth. Nonetheless, regardless of the existence of an active intention to deceive, they are still lying by definition.

Respecting the nuance of intent, we might adjust principle III to include this variation of conscious or subconscious intent, annexing the additional motive of lying to evade negative emotions like shame or discomfort. Additionally, principle II must be altered to respect the plurality of beliefs that may cause a lie. With this view, the necessary and sufficient conditions of lying can be restated as:

- I) *A* states that *p*
- &
- II) *A* believes that *p* is false by at least one personal belief
- &
- III) *A* intends to be deceptive to *B* in stating that *p*, meaning that *A*
 - i) consciously intends to deceive
 - or
 - ii) consciously intends to conceal information or evidence
 - or
 - iii) subconsciously intends to evade negative emotions.

By altering condition II, we include cases where, from the perspective of one of *A*’s beliefs, the belief expressed in *p* is false—even if it is true according to another conflicting belief. This accommodates cases like the religious father. From the perspective of his theological views, the belief that it is OK for his child to make the mistake of stealing is false, even if this is true according to his belief in parenting books and studies. But he may not be consciously aware of this disjunction, a problem that condition III addresses. By describing iii as a genuine form of deception, III now includes cases of self-deception, lies motivated by shame, compulsive lying, and other psychological impulses influencing unconscious deception. For the father, this manifests in his subconscious intent to evade the discomfort of his two contradictory, unresolvable beliefs. For the shame-evasive child, avoiding embarrassment is an unconscious motive. And for the compulsive liar, the agitation of denying an impulse—which is akin to ignoring a nagging itch—causes a split-second decision to lie and elude that feeling. For all three examples, this revised framework accurately assesses each additional motive for lying as valid.

I anticipate other philosophers responding to this proposal with a number of reasonable questions. Why should the condition of consciousness be reserved to i) and ii), while iii)

is the only type of deception regulated by the unconscious? If *A* has to believe *p* is false by only at least one personal belief, then does the definition of lying include too broad a scope of utterances? If the subconscious can contribute to a lie, then do people lie all or most of the time? While these are justified rebuttals, they can be addressed through counterexamples that lay bare the mechanics of self-deception for further examination. For example, the latter two questions may not necessarily be problems; lying *is* a relatively common human speech act, and a theory accounting for its frequency should not be discounted simply due to its expansiveness. However, to respond to the first query, applying the subconscious to part i) and ii) may in fact broaden the scope beyond what is helpful by creating inconsistencies in reasoning. The subconscious cannot be applied to part i) as intent necessarily requires a conscious notion of purpose. It would be equally difficult to incorporate the unconscious into part ii), as people rarely include every detail possible when answering broad questions like “How was your day?” By involving the unconscious with concealment, this may mean that anything one fails to mention about their day is a lie, which does not feel true in the practical sense. Ultimately, it might be necessary to distinguish self-deception from other forms of lying to avoid merging two kinds of non-alethic discourse; perhaps lying to oneself is an entirely different species than lying to others. An expansion of this theory must address and adapt to such critiques. However, I maintain that my psychological account provides direction to better demarcate the fluid borders of a lie. To refine the overly modest deceptionist view, philosophers must consider the variation in a liar’s intention and motives, as well as the influence of the subconscious.

Works Cited

Lackey, Jennifer. “Lies and Deception: An Unhappy Divorce.” *Analysis* 73, no. 2 (April 2013): 236–248. <https://doi.org/10.1093/analys/ant006>.

The Significance of the Fleeting: Happiness and Mortality in Augustine’s *De Trinitate*

Asal Arefi

Abstract: This paper challenges Augustine’s claim in *De Trinitate* XIII that the universal human desire for happiness entails a desire for immortality. Augustine argues that because mortal life is inevitably subject to loss, suffering, and death, it cannot satisfy the conditions required for genuine happiness, and that the will to happiness must therefore be oriented toward eternal life. I argue that this inference is unpersuasive, as it depends on the contested assumption that genuine value requires permanence. After reconstructing Augustine’s account of happiness as structured by the possession of what one wills, the right ordering of the will, and the absence of unwanted harms, I contend that his framework mischaracterizes the nature of human happiness. Many forms of human fulfillment are not merely compatible with finitude but are constituted by temporality, vulnerability, and the possibility of loss. Meaningful projects and experiences derive their value from unfolding over time, admitting risk and failure, and remaining subject to change. I further argue that Augustine commits a modal error by inferring from the desire for a good to continue within a temporal context to a desire for that good to exist eternally; wanting an experience to last longer does not entail a will for its eternal preservation, and in many cases its value depends on its impermanence. Finally, I suggest that finitude also renders suffering intelligible by situating it within a bounded temporal horizon, allowing happiness and suffering alike to be integrated into a coherent life. Absent the assumption that value requires permanence, the human desire for happiness does not entail a desire for immortality; mortality is instead a structural condition of meaningful human fulfillment.

“I leave Sisyphus at the foot of the mountain! One always finds one’s burden again. But Sisyphus teaches the higher fidelity that negates the gods and raises rocks. He too concludes that all is well. The universe henceforth without a master seems to him neither sterile nor futile. Each atom of that stone, each mineral flake of that night-filled mountain, in itself forms a world. The struggle itself toward the heights is enough to fill a man’s heart. One must

imagine Sisyphus happy.”

Albert Camus
The Myth of Sisyphus

Introduction

In *De Trinitate* XIII Augustine advances a striking claim about the structure of human desire: the universal human will for happiness, when properly understood, entails a will for immortality. Because mortal life is pervaded by contingency, loss, suffering, and death, Augustine argues that it cannot satisfy the conditions required for genuine happiness. Anyone who truly wills happiness, he maintains, must therefore will a form of life in which desired goods are secure and unwanted harms are excluded—a condition that only eternal life can provide. This argument plays a pivotal role in Augustine’s broader account of human fulfillment, linking ordinary practical desire to a metaphysical orientation toward permanence and incorruptibility.

This paper argues that Augustine’s inference from the desire for happiness to the necessity of immortality is ultimately unconvincing. While Augustine offers a sophisticated analysis of happiness as a state in which the will is rightly ordered, fulfilled, and unthreatened, his conclusion depends on a controversial assumption: that genuine value requires permanence. I contend that this assumption mischaracterizes the nature of human happiness. Far from being undermined by finitude, many of the goods that structure human flourishing—projects, pleasures, relationships, and achievements—derive their meaning precisely from their temporal, fragile, and vulnerable character. The possibility of loss is not merely compatible with happiness but often constitutive of its value.

By examining Augustine’s treatment of contingency, his rejection of Stoic and Epicurean strategies for accommodating finitude, and his appeal to the conditions of happiness articulated in *De Trinitate* XIII, I argue that his account conflates the desire for a good to continue within a temporal context with a desire for that good to exist eternally. This conflation, I suggest, amounts to a modal error: wanting happiness to endure does not entail willing it to be permanent, and the value of many human goods would be undermined rather than secured by eternal preservation. Once this assumption is rejected, the inference from the will to happiness to the will to immortality no longer follows. Mortality, rather than posing an obstacle to happiness, emerges as a structural condition of meaningful human fulfillment.

The Structure of Happiness in *De Trinitate* XIII

In Book XIII of *The Trinity*, Augustine begins with the widely accepted observation that all humans desire happiness, treating this not merely as a psychological tendency but as a structural feature of the will.¹ He then clarifies what it means to be truly happy by introducing three necessary conditions: [H1] one must have what one wills, [H2] one must not will

1 Augustine, *The Trinity (De Trinitate)*, 2nd ed., trans. Edmund Hill, O.P., ed. John E. Rotelle, O.S.A. (New City Press, 2015), 13.3–5; See also Augustine, *On the Happy Life (De beata vita)* 4.25, trans. G. E. M. Foley, modified; cited in Caleb Cohoe, “Philosophy as a Way of Life: New Research Directions,” in *Philosophy as a Way of Life: New Research Directions*, 30–55, DOI: <https://rb.gy/599elv>. Here Augustine emphasizes that a wise person cannot be made miserable by external events that occur against their will, since true happiness is grounded in virtue and the divine law of wisdom, which remain secure regardless of external circumstances.

what is morally wrong, and [H3] one must not have what one wills not to have.² Together, these conditions define what Augustine calls “living as he wills”:³ happiness is not simply possessing what one desires but also avoiding unwanted harms and rejecting what ought not to be willed. For the sake of argument, I accept these conditions as Augustine presents them.

At the same time, these conditions invite reflection and raise questions about their applicability to lived human experience. The first necessary condition [H1] suggests that happiness requires having exactly what one wills, but much of what we find fulfilling arises in the process of striving, learning, or engaging with challenges rather than in the mere possession of a desired end. Completing a creative project, mastering a skill, or building a relationship often provides joy in ways that cannot be reduced to achieving a static outcome; the process itself as the means to an end—the striving, engagement, and growth it entails—is central to what makes these experiences fulfilling independent of the ends themselves. Similarly, [H2]—the requirement not to will what is morally wrong—poses conceptual and practical difficulties. Morality is rarely absolute outside of a particular theological framework; desires we recognize as morally ambiguous or self-interested do not always render us unhappy, and moral growth can arise precisely from wrestling with and overcoming these desires. Even in moments when one wills something “wrong” in a conventional sense, the experience may contribute to self-understanding, character development, or the cultivation of empathy.

Condition [H3], perhaps the strictest of all, forbids experiencing anything one does not will. Yet much of what makes life rich—effort, risk, vulnerability, and even pain—would be disallowed under this standard. The intensity of a sporting competition, the emotional depth of a friendship, or the satisfaction of tackling a seemingly impossible project derives in part from the possibility of failure or loss. A mountaineer may relish the climb not despite the danger but because the very risk amplifies the accomplishment; a musician may feel exhilaration in performance precisely because it is fleeting and uncertain. In these cases, unwanted or “risky” elements do not undermine happiness but shape it. While I continue to accept the presented conditions, Augustine’s framework, while philosophically rigorous, sets conditions that exclude the very experiences that imbue mortal life with meaning, fulfillment, and happiness.

With these conditions in hand, Augustine turns to the human condition and the problem of contingency. Mortal life, he observes, is fragile and unpredictable: bodies deteriorate, relationships end, achievements fade, and death looms over all.⁴ Because we are exposed to harms we do not will, condition [H3] is constantly violated; even when we satisfy [H1] and [H2], the intrusion of unwanted events undermines the completeness of happiness.⁵

Philosophical attempts to adapt to this contingency, such as those of the Stoics and Epicureans, fail under Augustine’s framework. The Stoics argue that happiness can be achieved by aligning one’s desires with what is within one’s control and treating external goods as

2 Augustine, *Trin.* 13.5–13.7.10.

3 Augustine, *Trin.* 13.5.7, 13.7.10.

4 Augustine, *Trin.* 13.10–13; Augustine, *City of God*, cited in Diana Lobel, “Six St. Augustine: The Happy Life of the Soul,” in *Philosophies of Happiness: A Comparative Introduction to the Flourishing Life* (Columbia University Press, 2017), 781. Augustine here asserts that, given the mortal condition, all humans are inevitably subject to wretchedness, and that true happiness cannot be fully attained in this life, highlighting the limitations of temporal existence for human fulfillment.

5 Augustine, *Trin.* 13.7–10; See also Matthew Drever, “Augustine on Hope in Times of Suffering,” *Vox Patrum* 82 (2022): 145–166, 808, <https://doi.org/10.31743/vp.13048>. Drever explains Augustine’s point that genuine happiness is incompatible with fear of losing what one desires; for Augustine, the soul’s satisfaction depends on goods that are eternal and secure, which mortal life cannot fully provide.

“indifferents.”⁶ Similarly, the Epicureans locate happiness in bodily pleasure, which is fleeting but at least within one’s immediate experience.⁷ Augustine rejects both strategies: even if one endures suffering with virtue, unwanted harms still violate [H3]; and ephemeral pleasures cannot guarantee the possession of goods without threat of loss [H1] nor protect against misalignment of will [H2]. In short, neither adaptation nor pleasure secures the unthreatened good required for true happiness.

It is from these considerations that Augustine moves to his claim about immortality.⁸ Since genuine happiness requires satisfying all three conditions, and mortal life makes this impossible, the will-to-happiness must extend beyond temporal existence.⁹ Mortal life, however, is structurally incompatible with [H3]—the condition of “suffering nothing one wills not to”¹⁰—since vulnerability, loss, and death inevitably impose precisely the kinds of unwanted harms that the condition excludes.¹¹ In other words, anyone who truly seeks happiness must implicitly will a state in which unwanted harms cannot occur and desired goods cannot be lost—a condition that only eternal life can provide.¹² Augustine does not claim that

6 Augustine, *Trin.* 13.7.10; He notes that, according to Cicero’s account of the Stoics, a wise person achieves *apatheia*, a state free from emotional disturbances like fear, pain, desire, or pleasure; in this condition, external events no longer threaten happiness, highlighting the Stoic view that virtue alone suffices for a stable life, Teng He, “Is Virtue Self-Sufficient for Happiness?: Augustine on Virtue and Happiness,” *Journal of the Study on Religion and History* 1 (2025): 1–17, 4, <https://orcid.org/0000-0001-8350-3655>.

7 Lobel, “Six St. Augustine,” 159–160: “Augustine criticizes Epicureans for holding up pleasure as the goal of life—he mistakes Epicureans for sensualists—but he also criticizes the Stoics for taking pride in virtue, as if happiness is solely in human control. Augustine challenges the Epicureans: if we have a happy life, shouldn’t we want it to go on indefinitely?”

8 Avramenko emphasizes the contrast between temporal and eternal existence, noting that “time is time insofar as it contains change and motion; eternity is eternity only insofar as it contains no change and no motion.” He further highlights that human mortality is central to Augustine’s philosophy: “God is eternal, unmoving; man changes and is movable... This movement toward immanent death... lies at the core... of his thought, Richard Avramenko, “The Wound and Salve of Time: Augustine’s Politics of Human Happiness,” *The Review of Metaphysics* 60, no. 4 (Jun. 2007): 786–788, <https://www.jstor.org/stable/20130859>.

9 Augustine, *Trin.* 13.7–10; Augustine situates ultimate happiness in eternity, holding that mortal life alone cannot provide true fulfillment: as Lobel observes, he “believes eternal happiness is the only true happiness, and thus we must have the promise of an afterlife.” Yet temporal life is not devoid of meaningful experience, because hope allows the present life to participate in a foretaste of the ultimate good, Lobel, “Six St. Augustine,” 158–59; Drever notes that “hope does not leave us unhappy in this life, but neither does it reduce to a temporal form of happiness. Rather, hope gives us happiness now as a proleptic vision of Christ’s resurrection... toward true happiness,” suggesting that hope enables finite humans to apprehend a measure of fulfillment while still oriented toward the eternal. In this way, Augustine integrates the temporal and eternal dimensions of happiness, showing that human flourishing in time points beyond itself, Drever, “Augustine on Hope,” 158.

10 Augustine, *Trin.* 13.7.10.

11 Avramenko emphasizes Augustine’s focus on the temporal condition of human existence, noting that “God is eternal, unmoving; man changes and is movable... This movement toward immanent death... lies at the core... of his thought... to be human means to be timely, and to be timely means to be hurrying along the road to death,” underscoring how mortality fundamentally shapes the human experience of happiness. Avramenko, “The Wound and Salve of Time,” 787–88.

12 Lobel notes that Augustine “challenges the Epicureans: if we have a happy life, shouldn’t we want it to go on indefinitely?” highlighting his claim that mortal happiness points beyond temporal pleasure. Lobel, “Six St. Augustine,” 159; Drever observes that “And so all good and holy people, even amid torments of

people consciously articulate a desire for immortality; rather, the very structure of our desire for happiness, grounded in [H1]–[H3], reveals a deeper orientation toward permanence and incorruptibility.¹³ Thus, the human desire for happiness naturally points toward the will for eternal life.¹⁴

Augustine’s argument captures a subtle and enduring truth about human desire: our attraction to stability and security helps explain why fleeting pleasures often inspire a longing for permanence.¹⁵ Even if one ultimately questions whether eternal life is required for happiness, his framework illuminates the profound connection between the structure of human desire and the temporal, moral contours of life.¹⁶ Any critique of Augustine, therefore, must address whether this connection is necessary: does happiness truly require immortality, or can finite, mortal life provide a form of happiness that draws its value precisely from its impermanence?¹⁷

Temporal Happiness: Finitude, Value, and the Modal Fallacy in Augustine’s Argument

A major challenge to Augustine’s argument comes from a rival understanding of what happiness is. Augustine concludes that mortal life cannot satisfy human happiness, but this relies on the controversial assumption that genuine value requires permanence.¹⁸ If we adopt a different account of happiness, one rooted in the structure of finite human life, Augustine’s conclusion no longer holds.¹⁹

every sort, supported by God’s help, are called happy because of the hope for that end, the end in which they will be happy,” emphasizing Augustine’s view that hope frames temporal suffering within the expectation of eternal fulfillment. Drever, “Augustine on Hope,” 159.

13 Lobel notes that Augustine “believes eternal happiness is the only true happiness, and thus we must have the promise of an afterlife,” emphasizing the necessity of the afterlife for complete fulfillment; Drever complements this by observing that “Hope does not leave us unhappy in this life... toward true happiness,” showing that temporal happiness functions as a proleptic anticipation of the eternal, Lobel, “Six St. Augustine,” 159–160; Drever, “Augustine on Hope,” 158.

14 Lobel, “Six St. Augustine,” 159; Drever, “Augustine on Hope,” 158–159.

15 Lobel, “Six St. Augustine,” 159, notes that “Augustine challenges the Epicureans: if we have a happy life, shouldn’t we want it to go on indefinitely?” (On the Trinity 13.18.11), highlighting Augustine’s argument that true happiness points beyond temporal life.

16 Drever, “Augustine on Hope,” 158, observes that “Hope does not leave us unhappy in this life, but neither does it reduce to a temporal form of happiness. Rather, hope gives us happiness now as a proleptic vision of Christ’s resurrection... toward true happiness,” emphasizing Augustine’s view that hope allows temporal happiness to point beyond the finite toward eternal fulfillment.

17 Avramenko, “The Wound and Salve of Time,” 808, notes that “...one can never be truly happy if one lives in fear of losing that which satisfies one’s desires. One cannot live in fear, Augustine argues repeatedly, and be happy at the same time. The wisdom that satisfies the soul must be eternal and eternally one’s own,” highlighting Augustine’s argument that temporal insecurity prevents full happiness.

18 Avramenko, “The Wound and Salve of Time,” 780–81, observes that Augustine later sharpens this position in *City of God*, asking whether “man can have genuine felicity though mortal,” and concluding that “the more credible and probable position... is that all men, as long as they are mortal, must needs also be wretched.” This reinforces Augustine’s underlying assumption that genuine happiness requires permanence, since mortality itself is treated as sufficient to exclude felicity.

19 He, “Is Virtue Self-Sufficient for Happiness?,” 4, explains that for the Stoics, “happiness was characterized by permanence (*perpetuatio*) and constancy (*constantia*),” such that “pursuing bodily and external goods would affect the stability of happiness.” This conception illustrates the background assumption—shared by Augustine—that genuine happiness must be permanent, an assumption your argument challenges

On this alternative view, happiness is not defined by the possession of an *unlosable* good but by active engagement within a finite temporal life.²⁰ Human projects derive their meaning from precisely the features Augustine treats as defects: that they unfold in time, require sustained effort, admit the possibility of failure, and are shaped by risk.²¹ Finitude does not diminish value; it generates it. What is precious is often so *because* it can be lost.

Consider a simple example. If one is given a beautifully crafted gemstone, one may think its value lies in its intrinsic beauty alone, in the way it appeals to aesthetic sensibilities. Perhaps it could be admired regardless of loss or scarcity. Yet this reading is incomplete. While the gemstone's form and craftsmanship certainly matter, its significance is amplified by the fact that it is fragile and demands care. One guards it, cherishes it, and invests attention in it precisely because it might be lost. The possibility of loss is not a superficial add-on; it fundamentally shapes the way one relates to the object. Now imagine being told that identical gemstones appear automatically whenever one is misplaced or damaged. Suddenly, the gemstone retains its visual perfection but loses something essential: the lived experience of valuing, protecting, and attending to it. Its preciousness does not vanish entirely, but it is profoundly diminished. Beauty alone, divorced from context, engagement, and vulnerability, cannot carry the same depth of meaning.²² This phenomenon is not confined to gemstones. Human projects, relationships, achievements, and pleasures derive much of their significance from the tension between desire and uncertainty, effort and risk, presence and impermanence.²³ A musical performance captivates precisely because it unfolds in time, allowing mistakes, surprises, and ephemeral moments of brilliance. A relationship matters not simply because the other person exists, but because it requires investment, attention, negotiation, and mutual vulnerability. Even suffering, when bounded and meaningful, gains coherence through its temporality and

by grounding happiness in finite, temporal forms of life.

20 Avramenko, *The Wound and Salve of Time*, 808, summarizes Augustine's view that "...one can never be truly happy if one lives in fear of losing that which satisfies one's desires," since "the wisdom that satisfies the soul must be eternal and eternally one's own." This makes explicit the assumption your alternative account denies: that happiness must consist in the possession of an unlosable good rather than in finite, risk-laden engagement over time.

21 Avramenko, *The Wound and Salve of Time*, 786–788, emphasizes Augustine's sharp metaphysical contrast between temporality and eternity: "Time is time insofar as it contains change and motion; eternity is eternity only insofar as it contains no change and no motion," and accordingly, "God is eternal, unmoving; man changes and is movable—and specifically, man moves between life and death." This contrast clarifies why Augustine interprets change, risk, and failure as defects of the human condition, even as those very features render human projects intelligible and meaningful within a finite temporal life.

22 He, "Is Virtue Self-Sufficient for Happiness?," 4. Teng He explains that, for the Stoics, happiness is defined by "permanence (*perpetuatio*) and constancy (*constantia*)," such that "pursuing bodily and external goods would affect the stability of happiness." This framework illuminates the target of the gemstone example: by rejecting permanence as a criterion of value, the example directly undermines the Stoic—and by extension Augustinian—assumption that susceptibility to loss necessarily diminishes happiness, suggesting instead that fragility can be constitutive of meaning rather than a threat to it.

23 Drever, "Augustine on Hope," 158. Drever writes that "hope does not leave us unhappy in this life, but neither does it reduce to a temporal form of happiness. Rather, hope gives us happiness now as a proleptic vision...toward true happiness." This account supports the claim that meaningful human fulfillment operates within temporality rather than through immunity to loss: hope sustains happiness precisely amid vulnerability and suffering, indicating that significance arises from finitude and anticipation rather than from the possession of an unlosable good.

contrast with joy. To claim that beauty or value exists independently of risk or finitude is to miss the structural feature of human experience: things matter most when they are engaged with in the context of temporality, fragility, and effort. Far from being incidental, vulnerability and impermanence are constitutive of meaning; they are not a limitation to overcome, as Augustine suggests, but the very condition under which happiness and value arise.

Indeed, human experiences gain much of their significance from their fleeting, time-bound nature. A song we love delights in part because we cannot play it endlessly; repetition erodes anticipation, contrast, and the sense of discovery. The beauty of a sunset exists in its transience; extend it for decades, and awe dissolves. A photograph of a joyful moment preserves the form but not the substance: one does not feel the breeze, the laughter, or the lived experience. Similarly, a pinch of sugar enhances flavour precisely because it is limited—too much dulls the experience. The fleeting nature of these experiences, their impermanence, is what makes them meaningful and cherished.²⁴ Our desires themselves fluctuate—we may crave sugar at one moment and not the next, we may long for a song and later lose interest, or desire companionship at one time but retreat at another. Augustine's framework does not demand that desires remain static; one could, in principle, experience this kind of changing desire in an eternal life. On this reading, it may seem that mortality is not strictly necessary for meaningful engagement, since eternal life could, in theory, preserve dynamic desire.

Yet even granting this, the experiential richness of temporal life cannot be replicated by eternity. In a mortal context, the very alternation between wanting and not wanting, the anticipation of a sweet treat, and the knowledge that it will soon be gone all shape how we engage with and value the experience. These temporal boundaries create tension, surprise, and contrast, which give pleasures and projects their depth. In an eternal, unbounded context, even changing desires would unfold in a qualitatively different way: there would be no urgency, no finite horizon against which loss and gain acquire meaning, and the stakes that make ordinary joys vivid would flatten. Yes, the craving of a pinch of sugar may dissolve into indifference from one moment to the next even in eternal life—pleasure could respond to the ebb and flow of want. Yet it gains depth from temporality: knowing the treat is limited, anticipating it, and experiencing its fleeting sweetness intensifies the pleasure. The craving, the momentary indulgence, and the inevitable passing of desire are inseparable from the vividness of the experience. In eternity, the pinch of sugar might still be desired now and then, but without the temporal horizon that gives stakes and contrast, its significance would be muted. The fluctuating nature of desire might technically exist in eternity, but the meaning, intensity, and structure that finitude imposes on human life remain absent.

The temporal arc also enhances value through anticipation, reflection, and the contrast between presence and absence.²⁵ Longing for lost joys, remembering sunsets, or awaiting the

24 Avramenko, "The Wound and Salve of Time," 795; Lobel, "Six St. Augustine," 159–160. Avramenko emphasizes that time is experienced as difference and suffering, such that any account of happiness must contend with the unavoidable temporality of human life rather than its elimination. Lobel complements this by noting that Augustine, despite his emphasis on mortality and loss, nevertheless acknowledges the intrinsic goods of temporal life—virtue, achievement, and aesthetic beauty. Together, these observations reinforce the claim that the value of human experiences is bound to their temporal structure: meaning arises from contrast, anticipation, and transience, not from endless repetition or permanence.

25 Avramenko, "The Wound and Salve of Time," 795. Avramenko characterizes time as an experience of difference that is endured and suffered, such that human happiness must reckon with temporality rather than bypass it. While Augustine interprets this condition primarily as a deficiency to be overcome by eternity, Avramenko's analysis also underscores why anticipation, reflection, and contrast are structurally inseparable

next cherished song amplifies their meaning. Happiness is realized in the flow of experiences, not in a single, eternal state.²⁶ Were joy eternal and unchanging, humans would not even recognize its intensity; every moment would blur into sameness, stripping life of surprise, urgency, and significance. To live eternally in an unchanging state is, in effect, to lose life itself: the fleeting, dynamic quality that gives meaning to experience disappears. Eternity, rather than guaranteeing happiness, would flatten value: courage, hope, commitment, and desire would vanish, and the very structure that makes life meaningful would collapse.

If this is so, temporality is not an obstacle but a condition for meaningful engagement with life. Augustine's inference from the natural desire for happiness²⁷ to the necessity of eternal life²⁸ thus presupposes that a good must be permanent to be genuinely good.²⁹ Without this prior assumption, desiring happiness does not compel a desire for immortality.

One might concede Augustine a small victory: surely a runner in the midst of a euphoric high, or a person enjoying a trip or special moment, wishes the feeling could last “forever.” Augustine could argue that this demonstrates the human will's natural orientation toward permanence,³⁰ yet closer examination reveals a crucial distinction. The runner's desire is finite and contextual: she wants the sensation to continue a little longer, not to exist eternally in a frozen state. Human experience teaches that the fleetingness of happiness is essential to its value; permanence would render it banal or meaningless. Wanting a moment to continue does not imply wanting it endlessly. The inference from temporal to eternal is, at best, a conditional or modal error: a necessary condition for happiness now is mistaken for a necessary truth about desire forever.

This modal fallacy is evident in everyday examples. Consider a simple pleasure: a cup of coffee. In the moment, we may wish it not to spill, yet this does not imply a desire for infinite cups of coffee. The enjoyment of one cup is contingent on its temporality; if coffee were available eternally, each cup would lose its significance. Similarly, the runner may feel a euphoric high and wish it could last a few more minutes. But she does not wish to exist perpetually inside that state; the value of the experience comes precisely from its temporality. If

from temporal life: value emerges through difference, succession, and affective tension, all of which depend upon the very temporal arc Augustine treats as an obstacle to happiness.

26 Lobel, “Six St. Augustine,” 159–60. Lobel notes that “despite his catalog of the ills of our mortal life, Augustine also recognizes the wonders of this life, including our capacity for virtue, our achievements in the arts and sciences, and the aesthetic beauties of the world.”

27 Augustine, *Trin.* 13.5–8.

28 Augustine, *Trin.* 13.7–10.

29 Lobel, “Six St. Augustine,” 159–60. As Lobel explains, Augustine holds that “eternal happiness is the only true happiness, and thus we must have the promise of an afterlife...strictly speaking, we cannot [be happy in this life], although we can be called happy if we have the firm hope of eternal happiness.” This formulation makes explicit the assumption underwriting Augustine's inference: that a good counts as genuinely happiness-conferring only if it is permanent. Once this assumption is rejected, the move from the desire for happiness to the necessity of immortality no longer follows.

30 He, “Is Virtue Self-Sufficient for Happiness?,” 6. As He notes, Augustine holds that “happiness does not lie in the possession of external and bodily goods, but in the possession of the unchangeable good,” since virtue grounded in good will aims at what cannot be lost. This helps explain why Augustine would interpret the wish that a pleasurable experience last “forever” as evidence of the will's orientation toward permanence; however, it also clarifies the contested assumption at issue, namely that desiring a good's continuation commits one to valuing it only insofar as it is unchangeable, rather than temporally bounded and context-dependent.

it were eternal, the thrill would flatten into monotony. These experiences gain meaning from vulnerability and transience.

In each case, the desire is for a temporal continuation, not eternal preservation. Augustine might respond that even fleeting pleasures reveal a deeper orientation of the will toward permanence: the runner or the coffee drinker could be seen as implicitly desiring a state in which the good is unthreatened, echoing his condition [H3]. Yet this observation applies most clearly to ephemeral, consumable pleasures. What about intangibles that we hope might endure such as love and trust in a romantic partner, the bonds formed through friendship, or ideals like justice, peace, and knowledge? Surely, we often value these for their stability, and the threat of loss can seem corrosive.

However, even these goods are experienced and realized within temporal life, and their meaning depends crucially on the risk, effort, and attentiveness that mortal existence requires. Romantic love, for example, gains significance precisely because it demands commitment, communication, and vulnerability; the finite nature of life gives urgency and moral weight to acts of fidelity and care. Friendship deepens because we invest in it over time, navigating challenges, disagreements, and growth together. Ideals of justice, knowledge, and peace are structured by temporal struggle and uncertainty; pursuing these goods demands reflection, effort, and moral deliberation, and the possibility that they might be threatened or imperfectly realized gives ethical and experiential weight to our endeavors. It is precisely the process—the striving, maintenance, and care—that shapes our attachment and satisfaction.

In other words, the value of these enduring goods does not lie in their mere existence, but in the lived experience of pursuing, cultivating, and sustaining them. The process itself—tending to love, maintaining a friendship, or working toward justice and knowledge—engages us and creates emotional investment. We care about something precisely because it requires effort and responsiveness; we experience satisfaction not simply from the good existing, but in the labor of achieving or contributing to them. Finitude intensifies this: knowing that love could waver or that ideals could be compromised sharpens our awareness and imbues actions with moral and emotional weight.

Eternity, by contrast, would remove these stakes. If love, friendship, or justice were guaranteed forever, there would be no challenge to navigate, no vulnerability to manage, and no uncertainty to heighten appreciation. Achievements would feel hollow, relationships would lose their poignancy, and moral engagement would become rote. The work—the striving and care—is what produces the richness of attachment, satisfaction, and happiness itself, and the possibility of loss is what makes these experiences urgent, precious, and emotionally resonant. Finitude does not undermine happiness; it creates the conditions under which happiness can be fully realized, showing that mortality is not a limitation to joy or fulfillment, but the very structure that makes enduring human goods morally, emotionally, and experientially rich.

Augustine's reasoning thus overlooks the constitutive role of finitude:³¹ vulnerability, impermanence, and change are precisely what give our pursuits, attachments, and joys their meaningfulness. A mortal happiness that satisfies [H1]–[H3] need not exclude contingency or

31 Caleb Cohoe, “Augustine's Pursuit of Wisdom: Happiness as the Common Goal of Philosophy and Religion,” in *Philosophy as a Way of Life: New Research Directions*, ed. Caleb Cohoe (Brill, 2023), 43, https://doi.org/10.1163/9789004739147_004. As Cohoe observes, in our present life “our knowledge and love of God are sadly limited,” requiring “a more achievable standard for our thoughts and actions, especially for the knowledge that will guide us through our daily lives.”

risk; on the contrary, fully realized happiness in a temporal life necessarily includes engaging with these conditions. The ignorance of this, he mischaracterizes what it means to “fully” satisfy the will to happiness, making the leap to immortality unnecessary—and at worst, a modal error.

Finitude also shapes the intelligibility of suffering in a way Augustine does not fully consider. Just as happiness is fleeting, so too is pain, and this symmetry is crucial for understanding human experience. Our awareness that experiences—both positive and negative—allows us to place suffering within a meaningful temporal context. Knowing that moments of happiness will end naturally brings a degree of anxiety, but this same knowledge also implies that moments of misfortune, grief,³² or pain are similarly bounded. The awareness of transience moderates our fears and desires: we do not need eternal happiness to value it, because we understand that all experiences unfold within a finite arc. Therefore, the bounded nature of suffering is not a flaw to be eradicated by eternity (as H3 demands), but a structural feature that makes our temporal happiness coherent and complete.

Conclusion

Mortality, rather than being a deficiency to lament, functions as a structural condition that gives coherence and intelligibility to human life. The temporal arc of experience—its beginnings, middles, and ends—allows both joy and suffering to acquire significance. Happiness is appreciated not in spite of its impermanence, but because of it: pleasure is savored because it will pass, achievements are valued because they are not guaranteed, and effort carries weight precisely because failure is possible. Similarly, suffering is comprehensible only within a temporal framework; knowing that pain is bounded allows it to be endured and placed in perspective. The finitude of human experience, then, is not a barrier to fulfillment but an essential condition for the intelligibility of value, risk, and meaning.

This temporality directly challenges Augustine’s inference from the natural desire for happiness to the necessity of eternal life. While Augustine interprets fleeting pleasures as evidence of the will’s orientation toward permanence, closer analysis reveals that these desires are always conditional and context bound. A runner may wish a euphoric high to last a few more minutes, or a coffee drinker may wish their cup not to spill, but such wishes do not entail a desire for infinite repetition or eternal stasis. The modal fallacy arises when one assumes that because a momentary good is desired to continue, it must be desired to continue indefinitely. Human experience teaches that value is constituted by temporality; permanence would flatten intensity, stripping joy of its contrast with loss and rendering effort and achievement meaningless. The role of finitude also reframes the relationship between virtue, happiness, and human limitations. While Augustine critiques the Stoics for overstating the sufficiency of virtue, his own solution—orienting virtue toward God and eternal happiness—relies on an abstract, unachievable ideal. In contrast, recognizing happiness as temporally bounded allows virtue and effort to remain meaningful within the mortal sphere. Courage, prudence,

³² Grief, however, can stretch beyond immediate events, resurfacing long after loss, sometimes unpredictably. Yet even in these cases, the finite nature of life shapes how we experience and interpret it. Because our lives are bounded, we recognize that loss is part of a larger temporal arc: grief is not eternal, and our awareness that it will eventually ebb—or that life continues despite it—allows us to process, learn from, and integrate the experience into the broader narrative of our existence. The finitude of life gives us perspective, urgency, and opportunity for reflection, making even prolonged or recurring grief intelligible and, in a sense, meaningful. We can see it as part of a life lived fully, rather than an unending, formless state of suffering.

and temperance gain significance because humans must act amid uncertainty, vulnerability, and risk. Similarly, the fleeting nature of human knowledge and love does not diminish their value; rather, it makes the pursuit of understanding, attachment, and moral effort rich and consequential. The temporal framing of life enables a coherent, dynamic engagement with both pleasure and suffering without appealing to the necessity of eternal perfection.

Ultimately, the finitude of human life provides the very conditions under which fulfillment, meaning, and joy are possible. Mortality ensures that experiences are vibrant, that losses and gains are felt, and that human endeavors carry weight. It is within this temporal structure that happiness can be realized, not as a static or eternal state, but as a dynamic interplay of effort, desire, achievement, and reflection. Augustine’s insistence on eternal happiness overlooks the constitutive role of transience and contingency; the human will can be fully satisfied without aspiring to the immortal, unchanging good. In embracing finitude, we see that temporal happiness is not a provisional or incomplete good, but a deeply intelligible and experientially rich form of human flourishing.

Works Cited

- Augustine, Saint. *The Trinity (De Trinitate)*. 2nd ed. Translated and with introduction and notes by Edmund Hill, O.P. Edited by John E. Rotelle, O.S.A. New City Press, 2015.
- Avramenko, Richard. “The Wound and Salve of Time: Augustine’s Politics of Human Happiness.” *The Review of Metaphysics* 60, no. 4 (June 2007): 779–811.
- Cohoe, Caleb. “Augustine’s Pursuit of Wisdom: Happiness as the Common Goal of Philosophy and Religion.” In *Philosophy as a Way of Life: New Research Directions*, edited volume, 30–55. Brill, 2023. https://doi.org/10.1163/9789004739147_004.
- Drever, Matthew. “Augustine on Hope in Times of Suffering.” *Vox Patrum* 82 (2022): 145–166. <https://doi.org/10.31743/vp.13048>.
- He, Teng. “Is Virtue Self-Sufficient for Happiness?: Augustine on Virtue and Happiness.” *Journal of the Study on Religion and History* 1 (2025): 1–17.
- Lobel, Diana. “St. Augustine: The Happy Life of the Soul.” In *Philosophies of Happiness: A Comparative Introduction to the Flourishing Life*, 145–162. Columbia University Press, 2017.

Tainted Morality: Systemic Oppression and Moral Responsibility

Madelyn Stanley

Abstract: It is commonly believed that people who act autonomously are morally responsible for their actions. However, for those living under oppressive social conditions where their autonomy is challenged, whether from a lack of unforced options or a lack of moral epistemic background, the task of assigning moral responsibility for their actions is less obvious. On the one hand, it would be morally wrong to exempt people living under conditions of oppression from moral responsibility as this carries troubling consequences of stripping them of any claim to agency for their actions. On the other hand, it is difficult to assign appropriate blame when oppression interferes with autonomy. This essay explores moral frameworks and practices that consider how oppression can complicate and inform the standards for judging moral responsibility by answering the question: when autonomy is constrained by oppression, how should our practices of praise and blame adjust accordingly? I argue that while oppressive conditions may restrict agency, this does not eliminate moral responsibility for oppressed individuals, and I contend that analyzing oppressive contexts can offer valuable insight for improving our moral frameworks. I begin by constructing a theory of autonomy that preserves and sustains autonomy for individuals living under conditions of oppression. By demonstrating how autonomy is attainable for oppressed individuals, I then consider two theories that focus on the social aspect of moral responsibility and how we can adjust our moral practices according to one's social position.

Barry lives in a society that formally protects whistleblowers and values freedom of speech. When he witnesses an injustice at work, the risks Barry faces as a wealthy privileged natural citizen are minimal: perhaps some awkwardness with colleagues or mild social discomfort. Yet Barry chooses not to speak up. Maryam lives in that same society, but her position within it is very different. Maryam also witnesses an unjust act, and she too refuses to speak up. As a member of a marginalized immigrant community, she knows that speaking out against injustice can bring serious consequences: loss of employment, targeted harassment or retaliation against her family. Silence is the norm among those in her position. Should we blame Maryam less than Barry?

These examples highlight a tension between moral responsibility and the social conditions required for the exercise of personal autonomy. Moral philosophers have long noted that autonomy can be curtailed by oppression and social positionality, either because oppression limits one's ability to reflect critically, or constrains the range of options available.¹ While these opening examples are deliberately stark, oppressive conditions can be subtle and varied, though no less pervasive.² It seems reasonable to think oppressive conditions affect the degree of moral responsibility we assign, which raises the question: when autonomy is constrained by oppression, should our practices of praise and blame adjust accordingly? This paper explores how oppression complicates the standards for judging moral responsibility. I argue that while oppressive conditions may restrict autonomy, this does not eliminate moral responsibility for oppressed individuals. I contend that analyzing oppressive contexts can provide valuable insight for improving moral frameworks.

I begin this paper by illustrating how autonomy can be preserved under conditions of oppression using Diana Meyers' competency-based model of autonomy.³ Since we suppose that people who are responsible for their actions acted autonomously, in oppressive conditions where one's autonomy is seemingly threatened, Meyers instructively demonstrates how autonomy can be maintained under conditions of oppression. The second part of this paper explores how assigning moral responsibility is ultimately a social practice that should consider agents' actions and reasons from their social positionalities. Michael McKenna's conversational model notes that dialogical exchanges can reveal the meaning of a person's action and assign moral praise or blame accordingly. I contend, however, that McKenna's model for assigning moral responsibility falls short as he does not consider how conversational norms and biases can themselves be oppressive.

To repair these shortcomings, I turn to Manuel Vargas' scalar and forward-looking model of moral responsibility.⁴ He posits that oppression is a tool for moral accountability, improvement, and transformation. Meyers, McKenna and Vargas' accounts of autonomy and moral responsibility each play an essential role to develop this paper's broader methodological lesson: systemic oppression and injustice in our moral system can be confronted and remedied, if practices of moral praise and blame are conducted with attention to an agent's social positionality and their capacity to be autonomous, and recognize and respond to moral concerns.

The Impacts of Socialization on Autonomy and Moral Responsibility

Autonomy is central to moral responsibility because holding someone morally responsible presupposes their capacity to reason and act without external or psychological constraint. If an individual lacks the ability to make decisions independently, it is difficult to argue they

1 Christopher Bennett, "The Varieties of Retributive Experience," *The Philosophical Quarterly* 52, no. 207 (2002): 145–63; Victoria McGeer, "Mind-Making Practices: The Social Infrastructure of Self-Knowing Agency and Responsibility," *Philosophical Explorations* 18, no. 2 (2015): 259–81, <https://doi.org/10.1080/13869795.2015.1032331>; Susan Hurley, "The Public Ecology of Responsibility," in *Responsibility and Distributive Justice*, ed. Carl Knight and Zofia Stemplowska (Oxford University Press, 2013).

2 Iris Marion Young, "Five Faces of Oppression," in *Oppression, Privilege, and Resistance*, ed. Lisa Heldke and Peg O'Connor (McGraw-Hill, 2004).

3 Diana Meyers, "Personal Autonomy and the Paradox of Feminine Socialization," *The Journal of Philosophy* 84, no. 11 (1987): 619–28, <https://www.jstor.org/stable/2026764>.

4 Manuel R. Vargas, "The Social Constitution of Agency and Responsibility," in *Social Dimensions of Moral Responsibility*, ed. Katrina Hutchison, Catriona Mackenzie, and Marina Oshana (Oxford University Press, 2018).

can justly be held accountable for those decisions. Personal autonomy is traditionally understood as the capacity to govern oneself independently, rather than being governed by others. A common philosophical tenet is that in order to govern oneself, one must shape their life by making unimpeded decisions that are consistent with their will, desires and values.⁵ However, I argue that traditional accounts of autonomy focus excessively on merely intrapersonal conditions of autonomy and disregard the experiences of oppressed individuals, who are not independent from the wills of others.

It is against this background that I examine Diana Meyers' work where she posits that in cases of external social constraint, autonomy is still accessible. She illustrates how feminine socialization, the process by which girls are raised and encouraged to enact feminine gender norms, desires, and roles, may prevent women from meeting the criteria of autonomy in the traditional sense, which, in turn, contributes to their gender subordination. As women are socialized to cultivate feminine traits (e.g., empathy, accommodation, and care), and adopt feminine roles (e.g., wife, mother, domestic, submissive), they may be socialized to put others' needs before their own. A woman may feel obliged to sacrifice her desire for comfort to prioritize another's needs, which would undermine her capacity to act autonomously, because her decision, motivated by social expectations, does not align with her will and desire: society encourages her to sideline her desires for others' sake.⁶ Since these values, desires and roles are socially enforced, women's agency is bypassed, whether or not they genuinely endorse them.

Based on these social conditions, one may be tempted to infer that women do not meet the criteria for autonomy and may not be autonomous at all. That inference, however, carries a troubling consequence: it risks stripping women of their claim to agency in all contexts. In a sense, it punishes the victims of oppression by further dismissing their voices as non-genuine and not truly theirs. To counter this problem, Meyers contends that these socialized habits do not necessarily negate women's autonomy altogether, if an account of autonomy can consider social inequality. She proposes three alternative forms of autonomy that challenge the view of autonomy as an unconditional, absolute, or all-encompassing capacity: episodic, narrowly programmatic, and partial access autonomy. Each explains how autonomy is preserved despite feminine socialization.

Episodic autonomy denies that autonomy is a stable, ongoing state. Rather, autonomy can be exercised in discrete moments or 'episodes' when one makes a decision that aligns with one's immediate attitudes.⁷ For example, a stay-at-home mother may lead a prescribed life shaped by social expectations of fulfilling her domestic and mothering duties. Despite these prescribed expectations, she can make daily decisions where she exercises her autonomy in snapshot moments, such as choosing what to cook or who she interacts with. Episodic autonomy recognizes that people living under restrictive conditions may only find pockets of autonomy through occasional, immediate moments, which nonetheless matter for their self-expression and moral agency. In such cases, the fact that her agency is constrained in many domains by social structures does not remove her capacity to act autonomously with respect to these episodic opportunities.

The second form of autonomy, narrowly programmatic autonomy, is exercised when

5 Immanuel Kant, *Groundwork of the Metaphysics of Morals*, rev. ed. (Cambridge University Press, 2012).

6 Meyers, "Personal Autonomy," 624.

7 Meyers, "Personal Autonomy," 625.

someone makes small but consequential decisions about their life, even if their life, on a broader and more fundamental level, is not autonomous.⁸ For instance, a woman who has been socialized to want to have a family may not be able to decide whether she wants children, but she can make smaller-scale crucial, life-altering decisions about when to have children, how many to have or what contraception she uses. She can meaningfully assert her agency and make important life decisions even if they are partial and relatively narrow. While episodic autonomy addresses autonomous decisions that align with immediate attitudes in a particular moment, narrowly programmatic autonomy enables one to make more impactful decisions that can shape one's life in significant but comparatively small scope.

Lastly, partial access autonomy, describes autonomy that is expressed with varying degrees of intensity or completeness.⁹ Without fully or openly acting on one's desires or values, someone can modulate, conceal, or temper their autonomous expression due to social pressures or constraints. For example, one may downplay the strength of her desires or mask her true feelings to conform to social expectations yet still retain some degree of control over self-expression. Partial access autonomy recognizes that even limited or constrained expressions of autonomy can count as meaningful exercises of agency. Together, these three forms of autonomy, where autonomy may be discontinuous, domain-specific, or expressed partially rather than fully, expand the traditional conception of autonomy by accounting for the complexities of living under oppressive social conditions.

Meyers notably posits that these alternative forms of autonomy must be achieved through the exercise of a "repertoire" of cognitive skills, such as contemplation, evaluation, and imagination.¹⁰ This repertoire, which she calls "autonomy-competency," requires an active, deliberate control: "autonomy is something that a person *accomplishes*, not something that happens to a person."¹¹ To achieve autonomy-competency, one is propelled to engage these innate cognitive skills, all of which depend on one's social environment to enable the exercise of these skills. Meyers notes that some people may be unable to develop autonomy-competency due to oppressive social contexts that limit reflection, evaluation, imagination, or decision-making opportunities: for example, if someone never learned to contemplate their life or available decisions, Meyers believes they cannot develop autonomy.¹²

Meyers' analysis of autonomy-competency, however, replicates a difficulty she raised in the traditional accounts of autonomy in that autonomy can be curtailed by one's social circumstances: if one is not in an environment where they can develop this competency, they risk never enacting and developing the skills required for autonomy. This limitation undermines her account as she hopes to develop possibilities for autonomy yet posits that autonomy risks never being developed if these cognitive capacities for autonomy-competency are either present in an all-encompassing way or absent altogether. In response to this limitation, I attempt to preserve autonomy by offering a distinction between two types of social impact on autonomy. The first is the impact of social conditions on autonomy-competency itself (one's basic ability to exercise autonomy with one's skills), and the second is the impact of social conditions on the alternative accounts of autonomy we can access (episodic, narrowly programmatic, or partial forms). In the former, social conditions impact how an individual *becomes*

8 Meyers, "Personal Autonomy," 626.

9 Meyers, "Personal Autonomy," 626.

10 Meyers, "Personal Autonomy," 626.

11 Meyers, "Personal Autonomy," 626, emphasis added.

12 Meyers, "Personal Autonomy," 626.

autonomous (as in their ability to engage in their skills required for autonomy), whereas, in the latter type, social conditions impact how an individual *acts* autonomously (by engaging in her alternative forms of autonomy). One's *being* autonomous can be impeded to different degrees (mildly, severely, completely), and in different ways (affecting one's imagination, one's capacity to reason, etc.). Likewise, the ability to *act* autonomously can be impeded to different degrees and in different ways (how many episodic opportunities present themselves; how much partial agency can be expressed; how many programmatic choices are available). Complete denial of autonomy-competency may result in complete loss of autonomous action. But short of that, there are many combinations of these two aspects of autonomy that we could encounter. This distinction proves useful so as not to wholeheartedly abandon the 'all or nothing' possibility for autonomy: we may not *be* autonomous, but we can *act* autonomously sometimes.

Meyers' project denotes alternative forms of autonomy in light of socialization and oppression, relaying the complexities of achieving autonomy within the context of social inequalities and positionality, both of which are essential features of holding someone morally responsible for their actions. Now that I have demonstrated that autonomy is possible under conditions of oppression, the next part of this essay looks to McKenna and Vargas who illustrate how moral responsibility is a social and structural process that must consider oppression into the calculation of assigning praise or blame to moral actions.

McKenna's Conversational Model of Moral Responsibility

Michael McKenna considers the pertinence of the social aspect of moral responsibility, namely the relational and dialogical components. He believes we are responsible not just because of what we do, but how our actions are interpreted by and responded to within a moral community. McKenna's theory emphasizes a conversational model of moral responsibility. Moral responsibility, he claims, unfolds in a dynamic exchange between an agent and their community. The community evaluates the agent's "quality of will" as expressed through the explanation for their action and behaviour.¹³ If someone acts immorally and does not display a good quality of will, the community engages in a moral "conversation" with the agent, seeking an explanation to reveal their moral motivations and orientation. From this conversation, the moral community interprets the action and assigns praise or blame accordingly. This interpretive component is important: our actions do not transpire in isolation; they are assessed and interpreted by one's community. Through an exchange, the moral community assesses responsibility by getting a clearer picture of the agent's intentions and shapes their future behaviour by reinforcing moral norms.

McKenna highlights two social conditions that complicate the standards of holding people morally accountable in these conversations: control conditions and epistemic conditions.¹⁴ Control conditions concern the extent to which an agent is free to act according to their own will, without coercion or undue influence.¹⁵ For example, if Eric is ordered by his boss to fire a pregnant coworker for discriminatory reasons, and fears he might get fired if he

disobeys, then his moral culpability is mitigated by the coercive power his employer holds. Epistemic conditions, contrastingly, concern the extent to which an agent is reasonably expected to know or understand the moral weight of their actions given their social position.¹⁶ For example, if Carson, who was raised in an isolated, monocultural white community, compliments a racialized coworker's English skills, he may not understand his action for what it is: a microaggression. His compliment reflects his epistemic ignorance rather than his intentional malice: due to his unique interactions with fluent English-speakers, he is unaware how his comment reinforces harmful assumptions.

McKenna argues that control and epistemic conditions should be considered when assigning moral responsibility because they shape access to knowledge, freedom, and moral vocabulary. Epistemic and control conditions, however, do not absolve responsibility wholeheartedly; rather, it qualifies how we hold people accountable for their actions.¹⁷ We blame someone differently if they acted intentionally, or out of coercion or ignorance, because in the latter cases, we blame people as a means for them to learn and correct. In this way, McKenna considers a person's social identity and positionality in making a judgment of their moral conduct.

While McKenna highlights the control and epistemic conditions that influence whether a person deserves praise or blame for their actions, he does not consider how these conditions could affect the conversation that assesses moral responsibility. McKenna invites us to express ourselves to each other, yet he misses an important detail: just having the conversation is not enough. Dominant social groups often determine the moral standards that underwrite blame and praise. They control the very terms of the conversation—its norms, language, and etiquette. In other words, since responsibility is underwritten by conversational practices, control of conversational norms can entail control over who is heard, whose voices are authoritative, and which forms of speech are treated as legitimate expressions of agency. We must have conversations with an eye on the processes that make them unequal, and how oppression can impact interpersonal dialogue in the first place.

Miranda Fricker's concept of testimonial injustice is instructive to the concern for conversational power dynamics and biases as she highlights how marginalized individuals are often wronged in their capacity as knowers: their testimonies are dismissed, their moral perspectives are unheard, and experiences misinterpreted due to prejudice embedded in the moral community's conversational norms. Moreover, her concept of hermeneutical injustice highlights the absence of conceptual and interpretive frameworks that acknowledge and understand the experiences of marginalized groups. These experiences are rendered unintelligible to the collective that interprets moral actions.¹⁸ If actors' testimonies lack recognition and

16 McKenna, "Power, Social Inequities," 6.

17 Cheshire Calhoun, "Responsibility and Reproach," *Ethics* 99, no. 2 (1989): 389–406, <https://doi.org/10.1086/293071>; Michele M. Moody-Adams, "Culture, Responsibility, and Affected Ignorance," *Ethics* 104, no. 2 (1994): 291–309, <https://www.jstor.org/stable/2381578>; Jennifer Saul, "Implicit Bias, Stereotype Threat, and Women in Philosophy," in *Women in Philosophy: What Needs to Change*, ed. Fiona Jenkins and Katrina Hutchison (Oxford University Press, 2013), <https://doi.org/10.1093/acprof:oso/9780199325603.003.0003>; Jules Holroyd, "Two Ways of Socializing Moral Responsibility," in *Social Dimensions of Moral Responsibility*, ed. Katrina Hutchison, Catriona Mackenzie, and Marina Oshana (Oxford University Press, 2018); Michael Brownstein, "Attributionism and Moral Responsibility for Implicit Bias," *Review of Philosophy and Psychology* 7, no. 4 (2016): 765–86, <https://doi.org/10.1007/s13164-015-0287-7>.

18 Miranda Fricker, *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford University Press, 2007), 148.

13 Michael McKenna, "Power, Social Inequities, and the Conversational Theory of Moral Responsibility," in *Social Dimensions of Moral Responsibility*, ed. Katrina Hutchison, Catriona Mackenzie, and Marina Oshana (Oxford University Press, 2018), 5.

14 McKenna, "Power, Social Inequities," 10.

15 John Martin Fischer and Mark Ravizza, *Responsibility and Control: An Essay on Moral Responsibility* (Cambridge University Press, 1998), <https://doi.org/10.1017/CBO9780511814594>.

the hermeneutical resources to acknowledge and justify their actions, then our conversational practices of moral responsibility are biased and inadequate at thoroughly considering oppression. Fricker's work emphasizes the exclusion of marginalized people from the creation of moral standards, practices and conversations, which McKenna's conversational model overlooks. He does not acknowledge the social imbalances of moral meaning and discourse. This prompts an important consideration: if moral responsibility relies on biased social dialogue, are we able to fairly and effectively assess moral responsibility through conversations?

McKenna's framework is appealing insofar as it invites us to explain ourselves to each other in conversation, fostering an understanding between members of the moral community. However, insofar as the norms and resources that guide these conversations are shaped by oppressive practices that overlook marginalized testimonies, experiences and understandings, the dialogical model is a limiting tool in assigning moral responsibility. We must have conversation with an eye on the processes (e.g., social norms and biases) that make it unequal. Vargas offers resources for answering how to repair or re-shape responsibility practices in light of structural inequalities and oppression.

Vargas' Social Constitution of Agency and Moral Reform

Manuel Vargas argues that we can morally improve the norms of the social practice of moral responsibility, even within oppressive contexts. I began this essay discussing Meyers, who offered a theory that bridges social considerations, such as the context of oppression, with autonomy. Building on this, I then turned to McKenna's proposal, which advances the discussion of conversational moral responsibility, understanding it as a social process of negotiation with one's moral community. Now, Vargas's framework allows us to take a normative and corrective possibility for moral responsibility, offering a way to cultivate agency that is responsive to bias, oppression, and the possibility of reform. To this end, he proposes an agent-cultivation model: actors can cultivate a form of agency that is sensitive to moral considerations such as potential biases, the pervasive impact of oppression and power imbalances, and the possibility of reform. In effect, this amounts to an ameliorative project of consciousness-raising, aimed not merely at individuals but at the very practices through which responsibility is attributed and negotiated.

As we saw earlier in Meyers's proposal, Vargas makes claims to internal capacities and one's social environment to support his ameliorative project. Psychological capacities, according to Vargas, include cognitive and affective dispositions, like reasoning, empathy, understanding and self-reflection, that allow one to recognize, deliberate, and respond appropriately to moral conflict.¹⁹ These capacities tend to be relatively stable, and they are formed through reactive attitudes and feedback one receives from their social environment.²⁰ For example, if Harper sees her brother's being punished for stealing, she cognitively assesses this consequence and understands that she should not steal if she wants to avoid punishment. Harper develops this cognitive reasoning because she was exposed to a situation that prompted her to practice and apply it, exemplifying how psychological capacities are partly socially constituted. The second component of psychological capacities is how a moral community

19 Manuel R. Vargas, "The Social Constitution of Agency and Responsibility," in *Social Dimensions of Moral Responsibility*, ed. Katrina Hutchison, Catriona Mackenzie, and Marina Oshana (Oxford University Press, 2018), 2.

20 Vargas, "The Social Constitution of Agency," 8.

instills values and commitments that influence rational decision-making.²¹ If a community demonstrably values family, its members' rationality tends to reflect that value. One's moral community is also shaped by power structures. Since both psychological capacities and environment are socially situated and interact with one another, our moral agency is ecologically embedded and responds to the social-moral ecology surrounding us. The ecology, however, may well include oppression.²²

Despite oppression's pervasive effects, Vargas warns against excusing moral responsibility outright based on subjugation. Doing so risks undermining individuals' cognitive capacities and agency and ultimately challenges their integrity.²³ Oppression may limit but does not necessarily eliminate morally adequate options. Vargas proposes a scalar and asymmetrical system of moral culpability that accounts for oppression's impact on individuals' beliefs and deliberations, including internalized harmful norms and cultural scripts.²⁴ For instance, a woman expressing misogyny shaped by pervasive patriarchal norms may be partially excused due to internalized oppression, but excusing her entirely would prevent her from re-evaluating and overcoming these sexist beliefs. Thus, oppression can diminish culpability but does not absolve agency or responsibility altogether.

This scalar approach indicates that moral responsibility is graded in proportion to social positioning within power structures. So, two individuals committing the same harmful act may bear different levels of moral culpability depending on their social position. Vargas suggests that oppressors, whether consciously or not, can benefit from nefarious social arrangements, and may be responsible for the production or persistence of oppressive social circumstances.²⁵ Importantly, the powerful and subordinate are both held responsible, though, to different degrees and for different reasons.²⁶ The proportionate praise and blame affirm responsible agency even amid oppression, and this can prove to be a powerful tool to highlight overlooked moral considerations and opening avenues for cultivating moral sensitivity.²⁷ If oppression is factored into moral systems, we thereby become sensitive and conscious of oppression as a morally sensitive phenomenon that can harm someone. Developing this awareness can be a strategy for moral growth, which Vargas calls the agent cultivation model.

Vargas is attempting to resolve a dilemma with the agent cultivation model. On the one hand, he aims to give an account of autonomy that is sensitive to descriptive, non-ideal facts of one's social environment. On the other hand, he seeks a normative account, not one that merely describes existing practices, but one that can tell us what we ought to do, qualified by our current structures of oppression. He emphasizes the potential for moral agents to critically reassess and improve their moral beliefs and practices. Another component to the agent cultivation model is a relativized "ideal observer" who serves as a guiding morally exemplary figure who, within similar social and cognitive constraints, acts to best achieve the aims of a just moral responsibility system.²⁸ This "ideal observer" is not a transcendent or God-like figure who sets moral standards, rather one who provides a realistic aspirational

21 Vargas, "The Social Constitution of Agency," 9.

22 Vargas, "The Social Constitution of Agency," 10.

23 Vargas, "The Social Constitution of Agency," 17.

24 Vargas, "The Social Constitution of Agency," 18.

25 Vargas, "The Social Constitution of Agency," 18.

26 Vargas, "The Social Constitution of Agency," 12.

27 Vargas, "The Social Constitution of Agency," 17.

28 Vargas, "The Social Constitution of Agency," 14.

standard for agents to cultivate their actions and moral sensitivity. Using this ideal as a guide, individuals can reflect on their conduct and strive to improve, even while embedded in oppressive circumstances. The project of moral responsibility and the goal of praise and blame is to cultivate agency in individuals so they can appropriately respond to moral situations and develop their moral sensitivity.

One limitation of Vargas' model is that the ideal observer remains deeply influenced by their moral community, which is itself shaped by oppressive social structures. Even well-intentioned moral exemplars may unknowingly perpetuate systemic biases because they cannot transcend societal conditions. This poses a circularity problem in Vargas' framework. Complete escape from the influence of power, domination, and oppression is nearly impossible—even acts of resistance arise within these power dynamics.²⁹ However, oppression is always a factor, and moral reform necessarily occurs within an oppressive society. This is why including oppression within the bounds of morality and using oppression as a tool for moral consideration in hopes of improvement is a reconciliatory aspect of this social framework of moral responsibility. Actively considering ones and others' positionalities and access to agency within power dynamics when evaluating morality proves to be a helpful tool to evaluate the conditions of moral practices and systems.

It is important to also emphasize that the practicality of considering oppression should not be excessively leaned on as it could lead to the romanticization or glorification of conditions of oppression as a 'superpower' to mending social justice. Using experiences of the oppressed to tackle oppression gives responsibility to the oppressed to mend our moral systems rather than having the non-oppressed, powerful, and influential enact meaningful change to the moral systems themselves. While the non-oppressed should listen and understand the experiences of the oppressed and to provide more space to and uphold the voices of the oppressed, this is far from sufficient to make reparations to the moral system.

Conclusion

When assessing individual culpability within oppressive social contexts, it is crucial to recognize how systemic factors shape moral decision-making and opportunities for growth. Social conventions around morality and responsibility often fail to account for these influences. To account for autonomy within oppressive systems, Meyers provides an appealing account of episodic, narrowly programmatic, and partial access autonomy. Meyers' preservation of autonomy allows for the possibility of assigning moral responsibility to oppressed individuals. With Meyers' framework in mind, McKenna then argues that assigning moral accountability proceeds through engaging in social conversation with one's moral community. His conversational theory of moral responsibility falls short as an ameliorative method of moral systems because conversations can enact social biases and norms, undermining the narratives and experiences of marginalized people. Alternatively, Vargas offers a reformatory account of moral cultivation that seeks to repair moral injustice beyond oppression. He incorporates oppression into a scalar model that holds individuals accountable while recognizing limitations imposed by social power imbalances.

Vargas' agent cultivation model and the concept of the ideal observer allow moral reform to emerge as an attainable, self-motivated project grounded in social reality. However, Vargas faces a difficult tension at the heart of non-ideal theories that also attempt to be normative. His reliance on relativized ideal observers risks importing the very conditions of

oppression which those observers are relativized against. My aim in this paper has not been to argue that non-ideal theories which qualify their accounts of autonomy by reference to actual mechanisms of oppression are doomed to fail. Rather, I have sought to show that such theories must consciously grapple with a circularity problem. To avoid abstraction that would overlook the specific conditions under which oppressed people exercise autonomy, these theories must give us relativized ideals. Yet in doing so, they make themselves especially vulnerable to the possibility that those ideals may unwittingly reproduce unnoticed conditions of oppression. As we have seen with all three thinkers, the possibility of falling into practices that perpetuate oppression is not completely avoidable. The task, then, is to continually revisit and scrutinize these ideals. Hence, Vargas's consciousness-raising project must itself remain critical of the ways in which it might be subject to the very forces it seeks to resist.

Though vulnerable to bias and wider structural influences, the ideal observer approach invites critical awareness and ongoing moral improvement. Centring oppression within this framework not only acknowledges experiences of oppression, but also challenges systemic injustices embedded in our moral systems, making it a promising path toward more equitable moral responsibility. These frameworks allow us to better approach the analysis of moral responsibility as illustrated in the examples of Barry and Maryam that open this paper. I have highlighted the complexity of a theory of moral responsibility that incorporates unequal distributions of power and its influence on individual agency in society and our moral communities. These considerations enrich discussions on morality as we hope to improve our societal conditions by developing wider and more thorough ideas of who and what is right or wrong.

Works Cited

- Bennett, Christopher. "The Varieties of Retributive Experience." *The Philosophical Quarterly* 52, no. 207 (2002): 145–163.
- Brownstein, Michael. "Attributionism and Moral Responsibility for Implicit Bias." *Review of Philosophy and Psychology* 7, no. 4 (2016): 765–786. <https://doi.org/10.1007/s13164-015-0287-7>.
- Calhoun, Cheshire. "Responsibility and Reproach." *Ethics* 99, no. 2 (1989): 389–406. <https://doi.org/10.1086/293071>.
- Foucault, Michel. *The History of Sexuality, Volume I: An Introduction*. Pantheon Books, 1978.
- Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: An Essay on Moral Responsibility*. Cambridge University Press, 1998. <https://doi.org/10.1017/CBO9780511814594>.
- Fricker, Miranda. *Epistemic Injustice: Power and The Ethics of Knowing*. Oxford University Press, 2007.
- Holroyd, Jules. "Two Ways of Socializing Moral Responsibility." In *Social Dimensions of Moral Responsibility*, edited by Katrina Hutchison, Catriona Mackenzie and Marina Oshana. Oxford University Press, 2018.
- Hurley, Susan. "The Public Ecology of Responsibility." In *Responsibility and Distributive Justice*, edited by Carl Knight and Zofia Stemplowska. Oxford University Press, 2013.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Revised edition, Cambridge University Press, 2012.
- Mackenzie, Catriona, and Natalie Stoljar. "Introduction: Refiguring Autonomy." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edit-

29 Michel Foucault, *The History of Sexuality, Volume I: An Introduction* (Pantheon Books, 1978), 96.

- ed by Catriona Mackenzie and Natalie Stoljar. Oxford University Press, 2000.
- McGeer, Victoria. “Mind-Making Practices: The Social Infrastructure of Self-Knowing Agency and Responsibility.” *Philosophical Explorations* 18, no. 2 (2015): 259–281. <https://doi.org/10.1080/13869795.2015.1032331>.
- McKenna, Michael. “Power, Social Inequities, and the Conversational Theory of Moral Responsibility.” In *Social Dimensions of Moral Responsibility*, edited by Katrina Hutchison, Catriona Mackenzie, and Marina Oshana. Oxford University Press, 2018.
- Meyers, Diana. “Personal Autonomy and the Paradox of Feminine Socialization.” *The Journal of Philosophy* 84, no. 11 (1987): 619–628. <https://www.jstor.org/stable/2026764>.
- Moody-Adams, Michele M. “Culture, Responsibility, and Affected Ignorance.” *Ethics* 104, no. 2 (1994): 291–309. <https://www.jstor.org/stable/2381578>.
- Saul, Jennifer. “Implicit Bias, Stereotype Threat, and Women in Philosophy.” In *Women in Philosophy: What Needs to Change*, edited by Fiona Jenkins and Katrina Hutchison. Oxford University Press, 2013. <https://doi.org/10.1093/acprof:oso/9780199325603.003.0003>.
- Vargas, Manuel R. “The Social Constitution of Agency and Responsibility.” In *Social Dimensions of Moral Responsibility*, edited by Katrina Hutchison, Catriona Mackenzie, and Marina Oshana. Oxford University Press, 2018.
- Young, Iris Marion. “Five Faces of Oppression.” In *Oppression, Privilege, and Resistance*, edited by Lisa Heldke and Peg O’Connor. McGraw Hill, 2004.

A Challenge to the Moral Argument: The Problem of Animal Suffering

Adrian Haagen

Abstract: Within contemporary theology, the moral argument is one of the most frequently cited challenges to the naturalist. Andrew Ter Ern Loke offers a new formulation to the argument, framing divine nature as the necessary ontological standard for moral realism. This paper will offer an internal critique of his framework. Under the assumption of an objectively good God, a tension emerges given the pervasive nature of the problem of animal suffering. Suffering for the nonhuman animal is hardwired into the biological obligations of the natural world; under the assumption of a morally perfect personal creator, gratuitous predatory violence intuitively conflicts with our ethical expectations. The framework is then confronted with a dilemma: this deity either follows an entirely separate standard of ethic or directly violates the very nature that is meant to ground the objectivity. Available reconciliatory efforts for the problem are expansive, often utilizing the preexistent theodicies offered for humans such as arguments from free will and soul-making, with both missing the epistemic weight to extend the logic from human to nonhuman animal. This has led to a contemporary approach in neo-Cartesianism—the attribution of moral irrelevance to a lesser conscious status. Yet, the offered cases lack proportionate empirical grounding. Thus, as it stands, it appears a new explanation for the problem of animal suffering is needed to coherently posit the notion of theistic moral realism without an appeal to ignorance. This paper contends that until these tensions are resolved, the moral argument loses its persuasive force against the naturalist.

Introduction

The moral argument is one of the most often invoked arguments against naturalism within contemporary discussions, with C.S. Lewis and William Lane Craig popularizing and formulating variants of the position. The most widely referenced, presented by Craig, goes as follows:¹

1. If God does not exist, objective moral values do not exist.
2. Objective moral values do exist.

¹ William Lane Craig, *Reasonable Faith: Christian Truth and Apologetics*, 3rd ed. (Crossway, 2008).

3. Therefore, God exists.

More recently, Andrew Ter Ern Loke has offered a new formulation of the argument, with noticeable refinements that distinguish his from the previous variations.²

1. A number of objective moral truths exist.
2. These objective moral truths are either metaphysically grounded in an impersonal entity, a non-divine personal entity, or a divine personal entity i.e., God, or they are brute facts.
3. These objective moral truths cannot be metaphysically grounded in an impersonal entity.
4. These objective moral truths cannot be metaphysically grounded in a non-divine personal entity.
5. These objective moral truths can be metaphysically grounded in a divine personal entity i.e., God.
6. These objective moral truths are not brute facts.
7. Therefore, these objective moral truths are metaphysically grounded in a divine personal entity i.e., God (From 2 to 6).
8. Therefore, God exists (From 1 to 7).

While more expansive than Craig's formulation, Loke's logic does not stray far from Craig's, he simply untints the hidden premises, dodging the ambiguity often criticized within the original position. Accordingly, I will utilize Loke's formulation of the moral argument. To summarize, the general idea is that moral truths are inexplicable within a naturalistic worldview and demand some form of metaphysical grounding through a divine entity. Therefore, as moral truths *do* exist, a divine personal entity (e.g., God) exists.

I will offer an indirect critique of the moral argument, examining the ethics of the deity invoked. Assuming the premises, we can infer theistic moral realism: a single standard of morality embedded within the nature of God. As posited by Loke³:

“What is good (e.g., love) is a property of God, thus goodness is not a higher standard of God but is a property of God's nature.”

Given this, if a divine personal entity were to deviate from this absolute standard of morality, they would thus be betraying their own nature. I will make the case for an under-discussed form of the problem of evil to demonstrate this variance. However, it is most important to note that any example of such a conclusion would fulfill the same purpose.

Consider the examples: A lion subdues a zebra by puncturing its trachea with its jaws—a necessary process the lion will repeat throughout its lifecycle. A snake injects a venom into its prey, followed by hours of paralysis; a biologically necessary act. Assuming a creator, this system—one built upon mandatory, seemingly gratuitous suffering—was both orchestrated and commanded. However, to the naturalist, this commander does not resemble what we should expect of a morally perfect divine personal entity. As such, the explanatory weight offered as the single best explanation for objective moral values is notably unconvincing. I will evaluate nonhuman animal theodicies and debates on consciousness to assess the coherence of an all-good deity, ultimately demonstrating the problem of animal suffering remains unsolved, highlighting a serious contention for the moral argument.

² Andrew Ter Ern Loke, “A New Moral Argument for the Existence of God,” *International Journal for Philosophy of Religion* 26 (2022), <https://doi.org/10.1007/s11153-022-09842-1>.

³ Loke, “A New Moral Argument,” 34.

Free Will Theodicy

In the discussion of any form of problem of evil, it is most important to attend to the general theodicies emphasized within their respective discourse. While there is an important distinction made between particularly human and nonhuman animal suffering, it is necessary to attempt an extension of the logic from one to the other to emphasize this distinction. The free-will theodicy is one of the most often cited in terms of human suffering, originating from Augustine. The core claim is that suffering acts as a direct consequence of freedom of action as a world containing both libertarian free will and the absence of suffering is incoherent given the current physical state.⁴ Assuming the inability to suffer, then, situationally, pulling the trigger of a firearm may either shoot the bullet, or jam the weapon. The physical properties would shift depending upon the state most suitable to prevent suffering within a given scenario. The individual laws in place would thus arbitrarily shift as there is no causal relation that connects one's suffering to the firearm. Also consider whether the individual firing the weapon is really free given they have willed the infliction of suffering. Even under the assumption of an all-powerful deity, the logical difficulty would remain—the shooter cannot retain free will while simultaneously being incapable of inflicting suffering within a world that allows for the physiological ability to suffer. This raises a couple questions; the first is whether an individual's free will is more important than evil,⁵ the second is if—and how—Heaven escapes this problem. However, these are outside of the bounds of this paper.

While this theodicy may potentially account for human suffering, it does not coherently extend to nonhuman animals. For the animal, suffering is most often not the byproduct of will, but of a necessity hard-wired into their biology. Though nonhuman animals exhibit signs of agency—as will be discussed further—they do not appear to have much of an incentive or ability to rationalize and communicate the moral implications of predation (i.e., mauling one another limb-from-limb to survive). Moreover, these actions are largely instinctual rather than deliberate. Assuming nonhuman animals possess some form of moral cognition, the theodicy may provide an explanation on an incremental scale. Even so, this answer is insufficient; suffering is not a simple addition to the animal kingdom—it is deeply woven into its structure; a structure that, assuming an omniscient deity, seemingly could have been entirely devoid of predation as vegetarian species exist. Perhaps an unknown reason exists for an all-good God to both permit and command such a pervasive level of torment, but pleading skeptical theism is precarious in this circumstance. Consider the Garden of Eden; biblically referenced as a utopian state of existence. If goodness is contingent on predation and suffering, how could what is considered a pinnacle of creation exist without either? So, to then digress, predatory violence does not appear to be a direct result of free will, but derives from a function within a perceptually cruel natural order. As such, the problem of *gratuitous* animal suffering is completely left unanswered.

Soul-Making Theodicy

C.S. Lewis popularized what is often referenced as the soul-making theodicy, which frames suffering as theologically necessary—a must have for moral and spiritual maturity.⁶

⁴ Augustine, *On Free Choice of the Will*, trans. Thomas Williams (Hackett Publishing Company, 1993), 2.7. (Original work published ca. 388.)

⁵ Brandon Robshaw, “A fundamental flaw in the free will defence,” *International Journal for Philosophy of Religion* 98 (2025): 107–121, <https://doi.org/10.1007/s11153-025-09962-4>.

⁶ C. S. Lewis, *The Problem of Pain* (HarperSanFrancisco, 2001; original work published 1940).

The books of Romans and James support a similar position (ESV, 2016, Romans 5:3-5, James 1:2-4), illustrating that faith building and joy are direct consequences of suffering. Despite this, within contemporary theology, the theodicy is quick to be denied to the animal. This includes Lewis, who argues that the capacity for pain in animals ultimately serves no purpose.⁷

Yet, Trent Dougherty denies this, formulating an argument for animal soul-making in his book *The Problem of Animal Pain*. Dougherty posits that a class of animals, much like humans, go through theosis—the process of unifying with God (going as far as to note they will become like ‘talking animals’ in reference to Narnia). Amidst deification, they will become rational substances, forming attitudes, opinions, and satisfaction through retrospection of their lives.⁸ Now, while conceptually whimsical at first glance, Dougherty’s argument may have some weight. Assuming animals transcend into Heaven and garner the ability to articulate language, as well as obtain a fundamental understanding of their purpose within the natural order, the blow of animal pain may soften. Nevertheless, Dougherty would first need to demonstrate that this is both probable, as well as solves the problem. I will contend that (i) the biblical narrative does not align with his argument, and (ii) the problem is largely left unresolved assuming the premises.

(i) Dougherty’s argument, which he names the transcendental argument for animal deification, consists of 13 premises. Yet, for the sake of brevity, I will summarize the logical structure. Dougherty affirms that animals have sentience, are made in the image of God, and have moral standing. Given that God is perfectly loving, it follows that he would act lovingly and justly towards all living beings. To alleviate the suffering imposed on animals, there would need to be a source of reconciliation or compensation, with the only loving solution being its inclusion within a greater good through soul-making. As appeals to cosmic order fall flat, as well as the preexisting impossibility for soul-making, the only available recourse is *future* soul-making—requiring both animal resurrection and deification. Therefore, if God is just and all-loving, it follows that animals must be resurrected and deified.⁹

Let us first evaluate Dougherty’s second premise: “Animals are made in the image of God.” Dougherty asserts an extension of Imago Dei unto nonhuman animals using Ecclesiastes 3:18-20, Genesis 1:30, and Psalms 104: 27-30—each mentioning “life” (or in Hebrew, *nephesh*, נֶפֶשׁ) or “breath” (Ruach, רוּחַ). First, consider his cited Ecclesiastes verse:

“Surely the fate of human beings is like that of the animals; the same fate awaits them both: As one dies, so does the other. All have the same breath; humans have no advantage over the animals. Everything is meaningless. All go to the same place; all come from the dust, and to the dust all return. Who knows if the human spirit rises upward and if the spirit of the animal goes downward.”

The phrases “surely” and “who knows” are curious in regard to Dougherty’s position. These are openly uncertain phrases that purposefully emphasize the speaker’s ignorance—a stark contrast from a doctrinal affirmation of Imago Dei.

Genesis 1:30 posits that all beasts of the Earth possess the “breath of life.” Yet, as Phil Halper distinguishes, this breath—the *nephesh* or *ruach*—is falsely interpreted by Dough-

7 Lewis, *The Problem of Pain*, 123.

8 Trent Dougherty, *The Problem of Animal Pain: A Theodicy for All Creatures Great and Small* (Palgrave Macmillan, 2014), 3.

9 Dougherty, *The Problem of Animal Pain*, 145.

erty.¹⁰ Halper distinguishes this as a life force given to all things; one given at birth and taken at death, as evidenced through Ecclesiastes 12:7: “and the dust returns to the earth as it was, and the spirit returns to God who gave it.” Furthermore, the extension of Imago Dei is a profound contrast from Genesis 1:27-28—“So God created man in his own image... and have dominion over every living thing that moves on Earth,” establishing a deliberate divine distinction between man and the animal. 2 Peter 2:12 also transcribes a vivid description of animals as “creatures of instinct, born to be caught and destroyed”—another disparity from that of divine equivalence.

(ii) For the sake of argument, let us grant the premise that animals extend into Imago Dei. Be that as it may, noticeable challenges continue to carry a distinguishable weight. First, drawing from Phil Halper’s expanded problem of animal suffering, we must ask ourselves; assuming animals undergo theosis, what exactly are they supposed to interpret as so necessary to God’s plan?¹¹ Furthermore, under the umbrella of the old Earth non-literalist, what are the animals meant to glean after deification about their suffering during the vast period in which man did not exist? Does this not deem their Earthly existence, as well as any forgone suffering, largely if not wholly arbitrary?

Second, numerous biblical accounts also exist in which God commands animals to be slaughtered or heavily wounded—seemingly egregiously. Consider God’s commands to Joshua: “You shall hamstring their horses, and burn their chariots with fire” (Joshua 11:6, ESV). Hamstringing, the severing of the Achilles tendon, appears disproportionately extreme considering this was meant to serve as punishment to the kings fighting against the Israelites. Also consider the story of Achan; after admission of theft, not only was Achan himself stoned, but his children, oxen, donkeys, and sheep were killed, likely in a similar manner (Joshua 7:24–25, ESV). Furthermore, Job’s family and livestock were also burned alive and struck down by God—not as a punishment or form of retribution, but to test character (Job 1:13–19, ESV). The list goes further, yet the point has been illustrated. Given these examples, it is curious as to why God seems to consistently arbitrate either crippling or wholly egregious punishments while less extreme means were available, particularly given the victims were innocent of the relevant offenses.

My challenge to Dougherty is this: What exactly are the animals before man, as well as those egregiously punished by God, meant to rationalize as necessary in the deification process? If anything, the rational being will see an absurdity in demanding such a step to atone for the sin (or test) of another; presumably they would not provide an offer of open-armed forgiveness and enthusiastic growth. It is also important to mention that this logic extends to both man and animal—deification alone does not account for what seems to be arbitrary punishment. Thus, even when granting Dougherty’s premises, a coherent solution to the problem of animal suffering has still yet to be provided.

Neo-Cartesianism and Consciousness

As the immediate theodicies appear to fail concerning nonhuman animals, some philosophers appeal to animal automatism and neo-Cartesianism. René Descartes posits the notion that animals are automata—biological machines that fulfill a role, creating the illusion

10 Phil Halper and Kenneth Williford, “A Soul-Making Theodicy for Animals?,” *International Journal for Philosophy of Religion* 97, no. 1 (2024): 45–60, <https://doi.org/10.1007/s11153-024-09943-z>.

11 Phil Halper et al. “Neo-Cartesianism and the Expanded Problem of Animal Suffering.” *International Journal for Philosophy of Religion* 94, no. 2 (2023): 177–198. <https://doi.org/10.1007/s11153-023-09875-0>.

of experience for the observer.¹² Under this view, we could assume a mutilated animal would equate to a malfunctioning robot, however, this position is not taken seriously within contemporary discussion. Through the neo-Cartesian position, however, we find a less extreme version of automatism that still attributes moral irrelevance to the nonhuman animal, but accepts some form of sentience.¹³ Now, assuming either position, as animals would not possess moral value, the problem of animal suffering is thereby solved. So let us evaluate whether the position holds much persuasive or epistemic weight.

Multiple versions of the neo-Cartesian position exist, yet I will only mention two relevant variants as they often retain many of the same problematic inferences contradicted by the case I will illustrate within the coming paragraphs. For those seeking an even more in-depth analysis, Phil Halper offers an exhaustive critique of the alternative arguments.¹⁴

C.S. Lewis introduces the notion of a succession of perceptions—an event A occurs, then B, then C—with the formulation of experience ABC as the consequent.¹⁵ To consciously experience this chronological succession of events, one must be able to stand outside of each individual moment, recognizing the transition between them. Something must retain continuity—a self—to fully interpret the sensations into that of an experience. Lewis defends that this is *potentially* what the nonhuman animal lacks, leaving pain as something that occurs within the animal rather than something the animal is acutely aware of. Interestingly, Lewis concedes that while some higher-order animals may perceive event B, the lack of ability to situate B within a sequence eliminates their experiential continuity, denoting their suffering as morally insignificant.

A second variance of this position derives from Michael Murray, who has articulated an evidential argument for neo-Cartesianism, illustrating a neurological defense for the notion that morally relevant suffering does not extend beyond humans.¹⁶ Murray presents a three-fold hierarchy of pain: (i) nociceptive response to stimuli, (ii) a subjective experience of pain, and (iii) the awareness of the pain. He explains that humans possess higher order mental states dependent upon the integrity of the granular prefrontal cortex. When this region is damaged (e.g., prefrontal lobotomy patients), the cognitive and emotional experience of pain becomes disjointed. Murray understands the paradoxical relation often reported from these individuals, noting that while they are aware of the pain, there is nothing inherently unpleasant about the experience—a blind pain. Note that the pain itself persists, yet the agent lacks a negative emotional response, effectively abolishing any suffering. Murray posits that sentient animals' granular prefrontal cortex is either absent, underdeveloped, or functionally limited, providing reason to deny the third level within nonhuman animals. This is the defense Craig affirms, describing this as “a tremendous comfort” to animal lovers, rendering empathy for Bambi as fallacious.¹⁷

I believe this is a puzzling formulation. Assuming animals reside within this second

12 René Descartes, *Discourse on the Method and Meditations on First Philosophy*, ed. Donald A. Cress (Yale University Press, 1996), 26. (Original work published 1641).

13 Halper et al., “Neo-Cartesianism.”

14 Phil Halper et al., “Against Neo-Cartesianism: Neurofunctional Resilience and Animal Pain,” *Philosophical Psychology* 34, no. 4 (2021): 474–501, <https://doi.org/10.1080/09515089.2021.1914829>.

15 Lewis, *The Problem of Pain*.

16 Michael J. Murray, “Neo-Cartesianism,” in *Nature Red in Tooth and Claw* (Oxford University Press, 2008), 41–72, <https://doi.org/10.1093/acprof:oso/9780199237272.003.0003>.

17 William Lane Craig, “Does God Exist? The Craig–Law Debate,” *Reasonable Faith* (2011), <https://www.reasonablefaith.org/media/debates/does-god-exist-the-craig-law-debate>

order of pain awareness, I have to ask whether this is coherent? Remember, an experience of suffering is the precise issue. A subjective experience of suffering cannot exclude an awareness of pain as the very state of suffering is contingent upon an experience. Without experience, there is no pain, and consequently there is no suffering. I believe Murray's position is confused.

Now, the second issue here is this notion that only humans are self-conscious. I believe this is proportionately unsubstantiated given the nature of the claim. Murray points to the prefrontal cortex, a part of the body either entirely absent from nonhuman animals or underdeveloped within other primates. To give credit where credit is due, this part of the brain, for the human, is a core component to being aware of oneself as an object, as well as thinking about the act of thinking—metacognition. The conclusion that Murray deduces from this, though, is that every animal separate from the human is missing a morally relevant form of sentience. This is a leap that is difficult to disprove by itself, however, I can build an argument for the opposing position. The following section will provide a substantive case for nonhuman animal consciousness.

To assess our current interpretation of any form of conscious state, it is necessary to first define our terms. Due to the inherent mysterious and subjective nature of consciousness, it is nearly impossible to fit into a single definition. Philosophers such as Ned Block and Thomas Nagel noted this difficulty and have developed what I believe to be the most viable notions available.¹⁸ I will defend five proposed notions, each attributing different criteria to a conscious experience. The list goes as follows: (i) wakefulness, (ii) perception and responsiveness, (iii) access consciousness, (iv) phenomenal consciousness, and (v) self-consciousness. To preface, while an organism may not exhibit all five aspects mentioned, the more critical the criteria, the stronger the case to attribute a higher order (or morally relevant in this instance) form of awareness.

(i) Wakefulness is a given for nonhuman animals—making any of my own philosophical inquiry redundant. I will proceed assuming this notion applies.

(ii) Neurobiological evidence suggests that animals do exhibit perception and responsiveness. Take studies with rodents, for instance, which have associated elevated stress hormones such as corticosterone with chronic pain—indicating a neurological response to physical perception.¹⁹ This is not in contention, and as the criteria are fulfilled, I will proceed on the assumption that this notion also applies.

(iii) Access consciousness involves the internal availability of information for deliberate integration and reason. Consider Aesop's fable paradigm: borrowing from “The Crow and the Pitcher,” a team of scientists demonstrated that New Caledonian Crows, much like apes, understood water displacement.²⁰ (Jelbert, 2014). Once being given the task of displacing water within a narrow tube, the crows quickly discriminated between water and sand, heavy and light, and small and large—skills comparable to that of a 5-7-year-old. The crows exhibited

18 Colin Allen and Michael Trestman, “Animal Consciousness,” *Stanford Encyclopedia of Philosophy*, September 21, 2020, <https://plato.stanford.edu/entries/consciousness-animal/>.

19 Richard L. Spinieli et al., “Persistent Inflammatory Pain Is Linked with Anxiety-like Behaviors, Increased Blood Corticosterone, and Reduced Global DNA Methylation in the Rat Amygdala,” *Molecular Pain* 18 (2022): 1–12, <https://doi.org/10.1177/17448069221121307>.

20 Sarah A. Jelbert et al., “Using the Aesop's Fable Paradigm to Investigate Causal Understanding of Water Displacement by New Caledonian Crows,” *PLoS ONE* 9, no. 3 (2014): e92895, <https://doi.org/10.1371/journal.pone.0092895>.

similar cognitive ability when given the trap-tube task, using natural tools to push or pull food out of a perspex tube. These are deliberate and goal-directed actions that require an internal availability of information—meeting the conditions of access consciousness.

(iv) Phenomenal consciousness, as illustrated by Thomas Nagel, is described as the subjective experience of what it is like to *be* an individual organism.²¹ While the problem of other minds limits access to the experiences of others, external consciousness is a reasonable inference. Similar neurobiology and elaborate communication give reliable grounds to reasonably assert that other beings retain a self. While nonhuman animals may possess different behaviors and neurological mechanisms, if these indicators are reliable enough to verify external human minds, it is epistemically consistent to extend this trust to nonhuman animals. Rejecting the phenomenal consciousness of every single external species risks logic akin to solipsism—a logical possibility, yet unreasonably skeptical.

(v) Admittedly, self-consciousness is a loaded and difficult to define term. Definitions include:²² (a) awareness of one’s body as a physical object, (b) bodily self-awareness, and (c) mental self-awareness. I will cover each’s respective basis.

(a) Consider Gallup’s mirror test.²³ It is widely understood from the results of his trials that primates use mirrors to inspect their bodies (e.g., dyed markings, surgical implants, and unseen body parts).²⁴ The full scope of the psychological implications of Gallup’s mirror test remains somewhat ambiguous as it is difficult to establish whether the animals interpret the mirrored image as themselves (i.e., the image qua proxy) or merely as another object. This is further complicated by later studies in 2002, which showed that out of 163 chimpanzees tested, only 73 passed the mirror test.²⁵ However, this does not imply that those who failed the test do not have an agency. Age, social comfort, and prior mirror experience are all variables that likely contributed to the results. Nonetheless, the inference stands: at a minimum, primates are aware of their bodies as physical objects.

(b) The limitations of Gallup’s classical mirror test are largely a product of the time. Yet, contemporary scientists have more tools than Gallup, as we can now account for the internal mechanisms behind self-awareness—linking it to the secondary somatosensory cortex (SII).²⁶ In the cited experiment, a macaque monkey was conditioned and habituated to a full-body mirror, developing a stronger sense of self-recognition over time. Through seeing its reflection, multimodal visual and somatosensory activation was observed within the SII, which is entirely separate from the neurological reaction to seeing another primate for instance. This

21 Thomas Nagel. “What Is It Like to Be a Bat?” *The Philosophical Review* 83, no. 4 (1974): 435–450.

22 Allen and Trestman, *Animal Consciousness*.

23 Gordon G. Gallup Jr., “Chimpanzees: Self-Recognition,” *Science* 167, no. 3914 (1970): 86–87, <https://doi.org/10.1126/science.167.3914.86>.

24 J. D. Smith, W. E. Shields, and D. A. Washburn, “The Comparative Psychology of Uncertainty Monitoring and Metacognition,” *Behavioral and Brain Sciences* 26, no. 3 (2003): 317–339, <https://doi.org/10.1017/S0140525X03000086>.

25 Marc Bekoff, Colin Allen, and Gordon M. Burghardt, eds., *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (MIT Press, 2002), 337.

26 Rafael Bretas et al., “Neural Evidence of Mirror Self-Recognition in the Secondary Somatosensory Cortex of Macaque: Observations from a Single-Cell Recording Experiment and Implications for Consciousness,” *Brain Sciences* 11, no. 2 (2021): 157, <https://doi.org/10.3390/brainsci11020157>.

suggests not just full-body recognition, but an understanding that their body is unique—not simple conjoined limbs. Thus, select nonhuman animals exhibit signs of bodily self-awareness.

(c) Consider grooming in primates, a both hygienic and social behavior that reinforces group bonds and general social cohesion.²⁷ Larger groups of primates require reduced grooming time to accommodate increased social demands, and thus what follows is a drop in cohesive and stable grouping patterns, culminating in an eventual group fission. The takeaway from this is that it is reasonable to assume grooming is used as a tool for self-soothing.²⁸ It is also worth noting that stereotypes (i.e., repetitive, functionless behavioral patterns) are notoriously found in just about every notable large-brained mammalian species. Just as humans twirl their hair, horses crib-bite and weave, pigs chain-chew, and dogs slink away after a lecture—each of which is stress-alleviating, tension-easing behaviors suggesting a form of emotional regulation unexpected in species lacking self-awareness.

To recap, select species release stress hormones in response to physical distress, demonstrate cognitive capabilities comparable to those of young children, exhibit tension-easing behaviors, and display mirror recognition both behaviorally and neurologically. Reflecting on both neo-Cartesian positions, what of the behavior from these animals resembles a being with its sentient lights turned off? Consciousness is not a binary system—it is not either on or off, but a spectrum. So, while a bee’s sense of self may be different from that of a primate, it does not mean they are not at all sentient.²⁹ Thus, deeming an inequivalence enough to disregard moral worth is an unsubstantiated leap. Even if we may not have a definitive answer as to whether or not the phenomenal experience of a bat is equivalent to that of a human, we have left little room for speculation within select species. If we accept such criteria onto humans, it is epistemically inconsistent to not apply it to nonhuman animals. If one were to advocate for automatism or neo-Cartesianism, a heavy burden of proof would lie on them to discount the available evidence.

In terms of Lewis and Murray’s respective positions, as per the section above, modern research has shown that animals do not just appear to be in pain; they exhibit neurological responses in a form often nearly indistinguishable from that of humans. Furthermore, animals do not simply appear to perceive events in a linear succession; it has been demonstrated that select nonhuman animals exhibit context-driven emotional responses, requiring a degree of diachronic processing. These behaviors and neurological reactions would be inexplicable assuming Lewis’ hypothesis, and extremely unexpected given Murray’s hypothesis, especially given the studies that have shown select species can recognize their body as distinctly individual from any other form of being. Thus, the burden of proof has not been met given the proportion to the evidence, which ultimately explains the unpopularity of neo-Cartesianism within the field.³⁰

27 J. Lehmann, A. H. Korstjens, and R. I. M. Dunbar, “Group Size, Grooming and Social Cohesion in Primates,” *Animal Behaviour* 74, no. 6 (2007): 1617–1629, <https://doi.org/10.1016/j.anbehav.2006.10.025>.

28 Corrine K. Lutz, “Stereotypic Behavior in Nonhuman Primates as a Model for the Human Condition,” *ILAR Journal* 55, no. 2 (2014): 284–296, <https://doi.org/10.1093/ilar/ilu016>.

29 Anil Seth, “Explanatory Correlates of Consciousness: Theoretical and Computational Challenges,” *Cognitive Computation* 1, no. 1 (2009): 50–63, <https://doi.org/10.1007/s12559-009-9007-x>.

30 Matthew N. Zippel et al., “Animal Emotions and Consciousness: A Preliminary Assessment of Re-

Theological Implications

I have presented a case highlighting the difficulty in answering the problem of animal suffering. Yet, this is not simply a challenge for the moral argument, but also for the widely accepted notion of classical theism. The implications of affirming the contradictory coexistence of an all-good deity and animal suffering are not a simple threat to a single apologetic strategy, but core theological commitments of Judaism, Christianity, and Islam. Now the broader question becomes one of salvageability.

A potential response lies in skeptical theism, though it proves admittedly unsatisfying. Presuming the existence of a maximal being, it is an expectation that certain divine intentions are inaccessible through our cognitive understanding³¹—ones that may very well explain away the problem of animal suffering. Perhaps animals are not experiential beings; or if they are, they possess a soul, one in which that positions suffering as necessary to theosis. Yet, the solution is unattainable within the human’s cognitive limitations.

Theological safety granted, this move is not satisfying, nor epistemically responsible. An appeal to our cognitive limitations bears a resemblance to a form of blissful ignorance, which is a far cry from what is necessary to reasonably sustain ambiguous apologetics over naturalism. Also note: when we accept an argument from an epistemological hand wave, the lack of space for meaningful discourse eliminates any need for discussion as it becomes an impossible to disprove safety net, ending all discussions in a standstill. To digress, in regard to the classical theist, while the problem of animal suffering may not act as a sweeping defeater, it remains a significant threat that demands further development within contemporary discourse.

As for the moral argument, the formulation becomes restricted to those who presuppose the notion of an all-good deity, perhaps acting as a mere nail in the coffin. For those lacking the presupposition, as the natural world does not appear consistent with such a being’s existence, the persuasive power is weak. An appeal to morality in this state becomes a non-sequitur—an indistinguishable state from the God-of-the-gaps, following the formulation:

1. *x* exists (e.g., objective morality).
2. *x* lacks a clear explanation.
3. Therefore, God exists.

Conclusion

This paper has presented a critique of the inference to a perfectly moral deity, which, consequently, undermines the moral argument for the existence of God. William Lane Craig and Andrew Ter Ern Loke infer moral realism, with the nature of a divine personal entity serving as the ultimate standard. Yet, in light of the pandemic and the seemingly gratuitous suffering found within the animal kingdom, it appears the very deity they invoke either does not adhere to these standards, violates their own nature, and/or arbitrarily determines wrongness. As Loke illustrates:

“any moral imperfection would be contrary to the moral truths which God grounds.”³²

searchers’ Perceptions and Biases and Prospects for Future Progress,” *Royal Society Open Science* 11, no. 11 (2024): 241255, <https://doi.org/10.1098/rsos.241255>.

31 Stephen Wykstra, “The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of ‘Appearance,’” *International Journal for Philosophy of Religion* 16, no. 2 (1984): 73–93, <https://doi.org/10.1007/BF00136567>.

32 Loke, “A New Moral Argument.”

However, I believe moral imperfection does not begin to capture the depth of the problem of animal suffering. Remember, the problem is not simply that animals suffer—it is the absurd pervasiveness. Why must animals mutilate one another to survive? Attempts at reconciliation are largely speculative, with answers found in animal deification and the rejection of a morally relevant form of agency within the nonhuman animal. These conceptions largely present as ad hoc or disproportionately substantiated in reference to the preexisting evidence. Consequently, for the naturalist, the classical theistic deity increasingly resembles not an objective moral arbiter, but an indifferent creator. A creator who arbitrated a biological system wherein those it inhabits, simply by personal standards and demands are doomed to suffer without clear cause, and without choice. This world does not bear a resemblance to one orchestrated by an objective moral entity, thus arguing that this very divine personal entity is the most viable solution for moral realism does nothing in the form of persuasion. Consequently, the naturalist is not left with clarity, but with a deeper, justifiable skepticism.

Works Cited

- Allen, Colin, and Michael Trestman. “Animal Consciousness.” *Stanford Encyclopedia of Philosophy*. September 21, 2020. <https://plato.stanford.edu/entries/consciousness-animal/>.
- Augustine. *On Free Choice of the Will*. Translated by Thomas Williams. Indianapolis: Hackett Publishing Company, 1993. Originally written ca. 388.
- Bekoff, Marc, Colin Allen, and Gordon M. Burghardt, eds. *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. MIT Press, 2002.
- Bretas, Rafael, Miki Taoka, Sayaka Hihara, Axel Cleeremans, and Atsushi Iriki. “Neural Evidence of Mirror Self-Recognition in the Secondary Somatosensory Cortex of Macaque: Observations from a Single-Cell Recording Experiment and Implications for Consciousness.” *Brain Sciences* 11, no. 2 (2021): 157. <https://doi.org/10.3390/brainsci11020157>.
- Craig, William Lane. *Reasonable Faith: Christian Truth and Apologetics*. 3rd ed. Crossway, 2008.
- Craig, William Lane. “Does God Exist? The Craig–Law Debate.” *Reasonable Faith*. 2011. <https://www.reasonablefaith.org/media/debates/does-god-exist-the-craig-law-debate>.
- Descartes, René. *Discourse on the Method and Meditations on First Philosophy*. Edited by David Weissman. Yale University Press, 1996. Originally published 1641.
- Dougherty, Trent. *The Problem of Animal Pain: A Theodicy for All Creatures Great and Small*. Palgrave Macmillan, 2014.
- Gallup, Gordon G., Jr. “Chimpanzees: Self-Recognition.” *Science* 167, no. 3914 (1970): 86–87. <https://doi.org/10.1126/science.167.3914.86>.
- Halper, Perry, and Kenneth Williford. “A Soul-Making Theodicy for Animals?” *International Journal for Philosophy of Religion* 97, no. 1 (2024): 45–60. <https://doi.org/10.1007/s11153-024-09943-z>.
- Halper, Perry, Kenneth Williford, David Rudrauf, and Perry N. Fuchs. “Against Neo-Cartesianism: Neurofunctional Resilience and Animal Pain.” *Philosophical Psychology* 34, no. 4 (2021): 474–501. <https://doi.org/10.1080/09515089.2021.1914829>.
- Halper, Perry, Kenneth Williford, David Rudrauf, and Perry N. Fuchs. “Neo-Cartesianism and the Expanded Problem of Animal Suffering.” *International Journal for Philosophy*

- of Religion* 94, no. 2 (2023): 177–198. <https://doi.org/10.1007/s11153-023-09875-0>.
- Jelbert, Sarah A., Alex H. Taylor, Lucy G. Cheke, Nicola S. Clayton, and Russell D. Gray. “Using the Aesop’s Fable Paradigm to Investigate Causal Understanding of Water Displacement by New Caledonian Crows.” *PLoS ONE* 9, no. 3 (2014): e92895. <https://doi.org/10.1371/journal.pone.0092895>.
- Lehmann, J., A. H. Korstjens, and R. I. M. Dunbar. “Group Size, Grooming and Social Cohesion in Primates.” *Animal Behaviour* 74, no. 6 (2007): 1617–1629. <https://doi.org/10.1016/j.anbehav.2006.10.025>.
- Lewis, C. S. *The Problem of Pain*. HarperSanFrancisco, 2001. Originally published 1940.
- Loke, Andrew Ter Ern. “A New Moral Argument for the Existence of God.” *International Journal for Philosophy of Religion* (2022). <https://doi.org/10.1007/s11153-022-09842-1>.
- Lutz, Corrine K. “Stereotypic Behavior in Nonhuman Primates as a Model for the Human Condition.” *ILAR Journal* 55, no. 2 (2014): 284–296. <https://doi.org/10.1093/ilar/ihu016>.
- Murray, Michael J. “Neo-Cartesianism.” In *Nature Red in Tooth and Claw*, 41–72. Oxford University Press, 2008. <https://doi.org/10.1093/acprof:oso/9780199237272.003.0003>.
- Nagel, Thomas. “What Is It Like to Be a Bat?” *The Philosophical Review* 83, no. 4 (1974): 435–450.
- Robshaw, Brandon. “A Fundamental Flaw in the Free Will Defence.” *International Journal for Philosophy of Religion* 98 (2025): 107–121. <https://doi.org/10.1007/s11153-025-09962-4>.
- Seth, Anil. “Explanatory Correlates of Consciousness: Theoretical and Computational Challenges.” *Cognitive Computation* 1, no. 1 (2009): 50–63. <https://doi.org/10.1007/s12559-009-9007-x>.
- Smith, J. D., W. E. Shields, and D. A. Washburn. “The Comparative Psychology of Uncertainty Monitoring and Metacognition.” *Behavioral and Brain Sciences* 26, no. 3 (2003): 317–339. <https://doi.org/10.1017/S0140525X03000086>.
- Spinieli, Richard L., Rafael Alves Cazuzu, Amanda Juliana Sales, et al. “Persistent Inflammatory Pain Is Linked with Anxiety-like Behaviors, Increased Blood Corticosterone, and Reduced Global DNA Methylation in the Rat Amygdala.” *Molecular Pain* 18 (2022): 1–12. <https://doi.org/10.1177/17448069221121307>.
- The Holy Bible: English Standard Version*. Crossway Bibles, 2016.
- Aquinas, Thomas. *The Summa Theologica*. Vol. 1. Translated by Fathers of the English Dominican Province. Christian Classics, 1952. Originally written 1265–1274.
- Wykstra, Stephen J. “The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of ‘Appearance.’” *International Journal for Philosophy of Religion* 16, no. 2 (1984): 73–93. <https://doi.org/10.1007/BF00136567>.
- Zipple, Matthew N., Caleb Hazelwood, Mackenzie F. Webster, and Marcela E. Benítez. “Animal Emotions and Consciousness: A Preliminary Assessment of Researchers’ Perceptions and Biases and Prospects for Future Progress.” *Royal Society Open Science* 11, no. 11 (2024): 241255. <https://doi.org/10.1098/rsos.241255>.

Aesthetics as a Humanism

Christina Pan

Abstract: Philosophical treatments of suicide have largely proceeded along two canonical trajectories: ethical and existential. Ethical frameworks, grounded in normative assessment, often risk intensifying self-condemnation by situating suicidal desire within the register of moral failure. Existentialist accounts, by contrast, tend to aestheticize or romanticize self-destruction as a mode of authentic resolution, thereby introducing a different but equally problematic form of closure. Both frameworks may have structural vulnerabilities when applied to the lived phenomenology of suicidal crisis, where the demand is not for judgment or transcendence but for the reconstitution of meaning within immanent life. This paper proposes aesthetic philosophy as a third, comparatively underexplored framework through which to reconceptualize suicide. In particular, I draw upon phenomenologically oriented theories of everyday aesthetics that emphasize perceptual, embodied, and affective continuities of existence. These approaches reorient philosophical inquiry from normative justification to modes of attention, disclosing ways of remaining in contact with life that do not require prior demonstrations of life’s value or coherence.

“There is but one truly serious philosophical problem and that is suicide.” Albert Camus opens *The Myth of Sisyphus* with this claim. “No,” presumably, should be the obvious answer to suicide. Yet understanding *how* one arrives at this seemingly obvious “no” proves considerably difficult, particularly for individuals in suicidal crisis who find “yes” more compelling. Some philosophical traditions suggest we should not attempt to *reason* through the suicide problem at all. Nietzsche suggests that evaluating life’s value involves an inevitable circularity: the will to live is not a conclusion we reason toward, but a precondition that makes reasoning itself possible. To judge life, we must employ the very faculties and drives that life has given us.¹ The question, then, is not whether we can rationally prove life is worth living, but whether we can create or discover conditions that make living possible and desirable. This limitation does not undermine philosophical attempts to understand suicide. Rather, it situates philosophical reflection in its proper place, as a secondary attempt to articulate and interpret features of life that are more fundamental than the reasoning used to analyze them.

Typical philosophical approaches to suicide include ethical and existentialist frame-

¹ Friedrich Nietzsche, “Twilight of the Idols,” in *The Anti-Christ, Ecce Homo, Twilight of the Idols, and Other Writings*, ed. Aaron Ridley and Judith Norman, trans. Judith Norman (Cambridge University Press, 1997), §2.

works, which offer distinct conceptual resources but also may create structural vulnerabilities for someone in suicidal crisis. Ethical frameworks can inadvertently provide language for self-condemnation through normative assessment, while existentialist approaches may diminish agency through deterministic reduction or validate death as meaningful resolution. After examining these limitations, I explore how an aesthetic framework might offer a different route for understanding and ultimately dissuading an individual who is seriously considering suicide from committing suicide. Specifically, I focus on phenomenologically-oriented everyday aesthetics that emphasize immediate, present-tense sensory attention rather than theories of beauty or formal judgment. This aesthetic approach offers a mode of continuing contact with existence that does not require the individual to first establish that their existence meets standards of worthiness, authenticity, or meaning. We will explore the aesthetic approach to dealing with suicide in two distinct methodologies: one, the immediate aesthetic engagement that provides reasons to continue living by enabling continued contact with existence without justification, and second, the problem of suicide and aesthetics situated within a larger philosophical conversation that challenges the question of whether any *justification* needed for the existence question is needed altogether.

Ethical and existentialist frameworks have long dominated philosophical approaches to suicide, yet both traditions are better equipped to address the question of whether one should live than how one might continue to do so. I argue that aesthetic frameworks provide conceptual resources particularly well-suited to that question, resources that ethical and existentialist approaches have largely left unexplored. By suicidal crisis, I mean a condition in which an individual is seriously considering ending their life under acute psychological and moral distress. I identify distinctive vulnerabilities that ethical and existentialist frameworks may create for individuals in crisis, then explore what conceptual possibilities emerge when we turn to the aesthetic approach, which seeks to answer the “how” of existence rather than the “why.” Specifically, I explore a non-pharmaceutical approach to the problem of suicide, encouraging the individual in crisis to reflect phenomenologically on what makes life still worth holding onto. Aesthetic approaches to suicide remain relatively unexplored in philosophical literature compared to ethical and existentialist treatments, making this investigation both timely and necessary.

Suicide as a Philosophical Problem

A philosophical problem requires conceptual analysis of fundamental categories like meaning, value, and existence rather than empirical observation alone. Suicide qualifies as such a problem because it directly concerns life and death with absolute stakes, demanding that substantive philosophical approaches incorporate practical understanding into one’s engagement with existence.

Philosophical frameworks themselves may play a role in understanding or exacerbating suicidal crisis. Regardless of suicide’s immediate causes, from either overwhelming pain, anticipated loss, or other circumstances, individuals who attempt to end their lives inevitably draw upon conceptual resources to understand meaning, value, and existence. The philosophical frameworks available to someone shape how suffering becomes intelligible, whether continued existence appears possible, and what language becomes available for justifying life or death. From this problem, we must ask if there are possible frameworks for thinking about life, existence, meaning, and value that inadvertently may give people the language to answer “no” to all those reasons to keep existing.

Ethical Dimensions of Suicide

Ethical frameworks often draw upon the strongest reasons for living: purpose, meaning, relationships worth preserving, and goals worth pursuing. But ethics primarily works by establishing standards for value and significance, and this evaluative structure can cut both ways. When someone’s capacity for ethical self-assessment is compromised, the same frameworks that once sustained them can become weapons turned against the self.

Philosophers have traditionally condemned suicide on two foundational arguments: one deontological, one teleological. Kant treats self-destruction as a categorical violation; we have duties to ourselves as rational agents that cannot be dissolved by circumstance. Suicide wields one’s person merely as a means to escape suffering, violating the Formula of Humanity.² This constitutes a contradiction in which the rational will destroys the very condition of its own moral agency. Aristotle condemns suicide as cowardice, a failure to face life’s difficulties, and, more fundamentally, as a violation of our *telos*.³ Suicide truncates this project, abandoning the possibility of living well. Both frameworks presuppose an intact ethical subject in their analysis, as they treat the person contemplating suicide as a rational agent capable of bearing duties and pursuing excellence, which is precisely what makes suicide a moral violation rather than a non-moral event.

Other ethical traditions permit suicide under certain conditions. Stoics like Seneca defend the “rational exit” when circumstances make virtuous living impossible. If reason and dignity are lost, departure becomes *justified*. Utilitarian ethics can justify suicide when continued existence produces more suffering than flourishing, accounting for effects on oneself and others. These permissive frameworks appear to offer escape from the Kantian-Aristotelian prohibition. However, each framework that permits suicide does so by invoking ethical *evaluation* itself as grounds for self-destruction. This evaluative structure can create a vicious recursion for individuals in crisis.

In the Stoic position, the “rational exit” is contingent on the agent’s capacity to judge that virtuous living has become impossible. But this judgment is itself supposed to be an exercise of virtue, specifically of practical wisdom and rational assessment. The Stoic must be rational enough to recognize they can no longer be rational, virtuous enough to see they can no longer be virtuous. The rational exit becomes permissible precisely when suffering has eroded one’s capacity for rational virtue, yet adjudicating whether this threshold has been reached requires the very rational capacity that such suffering has compromised. This coherence requires a stable observing self that can step outside its own degradation to pronounce judgment on it, unfortunately not the case for most suicidal individuals. And even where such capacities remain intact, the framework assumes that life requires justification—that one must be able to give reasons for continuing to live that can withstand rational scrutiny. This paper contests such a premise.

It should be noted that each of these ethical traditions contains more nuanced resources—virtue ethics, for instance, features sophisticated accounts of impaired agency through the concept of *akrasia*, and Kantian ethics has been developed in ways that complicate the picture of rational self-duty. However, what is relevant for the present argument is not whether these frameworks are ultimately defensible, but how their evaluative vocabulary functions when internalized by individuals actively engaged in suicidal ideation.

² The Formula of Humanity: “So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means” (Kant 4:429).

³ *Telos* (τέλος) meaning “end,” “purpose,” “goal,” or “completion.”

Ethical frameworks provide the vocabulary for self-assessment: virtuous or vicious, authentic or inauthentic, net positive or net negative. This creates what I will call a recursive trap operating at multiple levels. In his analysis of Kurt Cobain's suicide note in *Expression and Survival*, Craig Greenman identifies how ethical vocabulary saturates Cobain's self-assessment: calling himself a "moody baby," confessing to feeling "emasculated," and apologizing for his inadequacy. Cobain describes himself as too sensitive, unable to feel the music anymore, "ripping people off" by continuing to perform. The note reads as though Cobain is measuring himself against standards of authenticity, passion, and integrity that he could no longer meet. As Greenman observes, "the language of the suicide note is consistent with the ethical argument against suicide."⁴ Normative ethical theory may have provided the very measuring sticks against which Cobain found himself fatally deficient.

The recursive trap operates as follows. Ethics provides the vocabulary for self-assessment, then demands we take this self-assessment seriously as a guide to action. Finally, ethics insists we act with integrity, aligning our actions with our judgments. The synthesis becomes lethal; if I judge myself ethically deficient, and I take ethics seriously, and I value integrity, then continuing to exist in my deficient state becomes itself an ethical failure. The more rigorously one applies ethical reasoning, the more compelling the conclusion for life or against life becomes. Ethics, designed to guide human success, provides the logical language for assessing value that, when not met, can help an individual conclude that one should cease to exist.

The Existentialist on Suicide

Existentialist approaches initially appear to avoid the recursive trap that can ensnare ethical frameworks. Rather than measuring the self against normative standards of virtue or authenticity, existentialism stems from radical freedom. Sartre declares that we are "condemned to be free," thrust into existence without predetermined essence or purpose, bearing total responsibility for what we make of ourselves. "Man is nothing else but what he makes of himself," Sartre writes in *Existentialism Is a Humanism*.⁵ This framework appears promising because it does not tell the suicidal person they are morally failing or merely sick. It acknowledges the genuine philosophical weight of the question of suicide without imposing external standards of worthiness. But if human existence is characterized by radical freedom and the burden of creating one's own meaning, then suicide becomes one legitimate expression of that freedom. Sartre's insistence that "existence precedes essence" means we define ourselves through our choices, and suicide itself is a choice.

It should be acknowledged that existentialism is not monolithic. Some readings of Sartre emphasize responsibility to others and commitment to ongoing projects as constraints on radical freedom, and other existentialist thinkers resist the conclusion that all choices are equally authentic. However, the structurally permissive tendency of the tradition, particularly as it has been popularly absorbed, remains a live concern for the present argument. The general trend of the existentialist framework treats self-destruction as a radical authentic choice, where choosing death is as authentic as choosing life. At minimum, the framework risks neutrality between continued existence and self-destruction in ways that may be dangerous for those in crisis.

4 Craig M. Greenman, "Cobain's Suicide Note as a 'Last Work': Art, Authenticity, and Ethics," in *Expression and Survival: Studies in the Aesthetics of Extreme Experience*, ed. Jonathan Mitchell (Routledge, 2014), 27.

5 Jean-Paul Sartre, *Existentialism Is a Humanism*, trans. Philip Mairet (Methuen, 1948), 22.

Jean Améry, an existentialist philosopher who survived Auschwitz, articulates what may be existentialism's most troubling philosophically justified endpoint in his *On Suicide: A Discourse on Voluntary Death*. Améry argues that "voluntary death is a privilege of the human."⁶ An individual has the right to take complete possession of their life by defining its end. To Améry, suicide is essentially the ultimate self-determination in a universe where existence remains largely subject to external forces. Améry himself took his own life two years after publishing his discourse on suicide. While we might validate Améry's suicide as an authentic expression of existentialist freedom, his case exposes a structural risk within the existentialist treatment of suicide. Existentialism treats suicide as the ultimate authentic choice and, by doing so, risks elevating it as the pinnacle of human existence. For someone in suicidal crisis seeking reasons to live, existentialism may instead provide philosophical validation for death as a meaningful resolution.

This risk, visible in Améry's case, stems directly from existentialism's foundational commitments. Sartre's claim that we are "condemned to be free" means we bear total responsibility for what we make of ourselves. For someone already experiencing life as meaningless or unbearable, this may intensify rather than alleviate crisis. It suggests that their suffering is something they have chosen, something they are responsible for, and something only to be resolved by their own proper exercise of freedom. And when the burden of freedom becomes unbearable, existentialism presents suicide as one more authentic choice, as the ultimate exercise of freedom, the entire refusal of the burden.

Even Camus, who explicitly argues against suicide in *The Myth of Sisyphus*, offers limited resources for someone in crisis. Camus claims that recognizing life's absurdity while refusing suicide represents authentic human dignity. He concludes: "One must imagine Sisyphus happy."⁷ This phrase is memorable but opaque, as it tells us *what* we should do (imagine Sisyphus happy, rebel against absurdity) without providing resources for *how* to do so if one does not already feel that rebellious vitality. For our purposes of understanding which philosophical frameworks can help individuals in active crisis, absurdism may not even register. His strongly rejects suicide but provides little conceptual ground for someone who does not already feel that pull towards life.

The Aesthetic Approach

The ethical and existentialist frameworks that have dominated philosophical responses share a common premise, that evaluative judgment about life's worth is both possible and necessary for resolving the crisis. For someone experiencing suicidal crisis, however, this presupposition may be precisely what has collapsed, leaving them trapped in frameworks that demand the very capacities the crisis has rendered inoperative.

I propose a phenomenologically-oriented aesthetic approach to the philosophical problem of suicide meant to specifically address individuals in active crisis, one that sidesteps both the moralistic prohibition ("don't do it, life is valuable") and existentialist neutrality ("it's your authentic choice"). Aesthetics is a vast philosophical domain encompassing everything from Kantian formalism to evaluative theories of beauty, sublimity, and artistic creation. The framework proposed here lies outside the typical artistic analysis of aesthetics. The everyday

6 Jean Améry, *On Suicide: A Discourse on Voluntary Death*, trans. John D. Barlow (Indiana University Press, 1999), 43. Originally published 1976.

7 Albert Camus, *The Myth of Sisyphus and Other Essays*, trans. Justin O'Brien (Vintage Books, 1991), 123. Originally published 1942.

aesthetics I describe to reckon with the problem of suicide emphasize immediate sensory attention. This approach does not escape the evaluative features of the previous frameworks entirely but relocates the question. Rather than asking “on what basis is life worth living?”, it invites attention to present moments like textures, sensations, and details. The former framework provokes a global judgement on life, while the latter requires only local attention. In a way, aesthetic engagement constantly appraises surroundings (“this moment matters,” “this sensation registers,” “this detail commands attention”) but does not require that life as a whole be found worthy, meaningful, or authentic for one to continue living. This framework offers a third path by neither encouraging life nor remaining neutral about death, but instead providing resources for *remaining in contact with existence* without first resolving whether that existence is justified.

It is crucial to distinguish this proposal from a superficially similar but fundamentally different approach, the claim that one should notice beauty as evidence that life is worth living. For example, take what could be a common line from a well-meaning but misguided person trying to help someone in suicidal crisis: “life is so beautiful; don’t kill yourself.” Trapped within evaluative frameworks that make a global judgement about the state of life, this approach provides a reason to stick with it. By asking that one assess whether the presence of beauty meets a standard sufficient to warrant continued existence, it still presupposes the precise evaluative capacities that may have collapsed in crisis.

Previous philosophical approaches to suicide largely seek to determine whether suicide is rational, permissible, or authentic. But for someone in an active crisis, these neglect the more urgent question: not whether life is worth living in the abstract, but *how to remain in contact with existence* when the capacity to justify that existence has collapsed. Recent work in everyday aesthetics establishes that aesthetic experience operates fundamentally differently from rational justification in ways that make it accessible even when other capacities have been compromised. Yuriko Saito’s development of the field in *Everyday Aesthetics* demonstrates that aesthetic experience pervades ordinary life in ways that require no special training, talent, or accomplishment. Saito distinguishes between art and non-art, noting the aesthetic experience of many non-art experiences as universal; “Most non-art objects and activities concern our everyday experiences of eating, clothing, dwelling, cleaning, and dealing with natural elements.”⁸ A special mode of appreciation can come of noticing details like the texture of morning light falling across a table, the rhythm of walking down a familiar street, the taste of coffee, or the sound of rain against windows. These experiences demand nothing except attention to the qualitative dimension of ongoing experience as it unfolds moment by moment through whatever senses are available. One need not be talented, creative, or accomplished to notice these textures of experience. The specific modalities available vary for each person. Someone who cannot see might attend to textures or sounds; someone who cannot walk might notice the rhythm of breathing or the weight of their body in stillness. Regardless of sensory capacity, aesthetic engagement is accessible. For someone in crisis whose evaluative capacities have collapsed but whose sensory mechanisms persist, this accessibility proves crucial; aesthetic attention asks neither that the moment possess special beauty or significance, nor that one justify why this moment, rather than some other, deserves notice. Saito’s framework thus eliminates the requirement of achievement or expertise that might otherwise constitute a barrier to engagement.

This minimal requirement of simple attention to sensory experience might nevertheless

8 Yuriko Saito, *Everyday Aesthetics* (Oxford University Press, 2007), 88.

appear beyond reach for someone in suicidal crisis, whose capacity for sustained attention or coherent experience may be severely compromised. Sheri Irvin’s work in “The Pervasiveness of the Aesthetic in Ordinary Experience” addresses this concern by challenging the assumption that aesthetic experience requires unity, closure, and sustained conscious attention. Irvin argues against philosophers like Dewey who impose such requirements, demonstrating instead that aesthetic experience can be “simple, lacking in unity or closure, and characterized by limited or fragmented awareness” without losing its aesthetic character.⁹ She describes experiences that exist in what she terms the “twilight of consciousness,” moments in which “one is vaguely aware,” but not “vividly present.”¹⁰ The vagueness of awareness itself can constitute part of the aesthetic texture of a moment. Most significantly, these experiences need not build toward any resolution or culmination. Aesthetic attention can simply “drift away”¹¹ without reaching conclusion, without serving any larger purpose, and without accumulating toward judgment. One may notice the weight of a coffee cup, the texture of light on a table, or the sound of rain, all self-fulfilling moments. For someone whose attention fragments and drifts, Irvin’s framework does not disqualify aesthetic engagement. Even such a vague, drifting, fragmentary awareness retains aesthetic validity.

Granted one’s capacity to notice details without talent or sustained attention, why should one bother? What justifies attending to light or texture when one cannot establish meaning therein or come to conclusions about whether existence is justified? Martin Seel’s concept of *Erscheinen* (appearing) addresses this obstacle by revealing that aesthetic attention operates entirely outside the framework of justification. Seel differentiates between an object’s “being-so,” its fixed and conceptually determinable properties, and its “appearing,” the indeterminate process of how it presents itself in perception. Aesthetic perception, Seel argues, attends not to what an object is but to “the how of their givenness here,”¹² focusing on “the phenomenal simultaneity of the aspects sensitively perceivable in an object.”¹³ Appearing is foregrounded when we “allow an object of perception to have effect without restricting ourselves to specific aspects of its constitution or function.”¹⁴ The light appears; perception of it follows. No evaluative judgment mediates this engagement. Seel designates this mode “mere appearing,” (*Erscheinen*) where “we pay attention to nothing other than the repleteness of its momentary and simultaneous givenness”¹⁵ with “no ambition to constitute a meaning of what is sensed.”¹⁶ The appearing itself constitutes the engagement, requiring no intermediate evaluative step. For someone in suicidal crisis, this proves crucial because the collapse often occurs precisely in evaluative capacity: the ability to judge whether something matters, whether existence is justified, and whether this moment deserves attention. Seel’s framework reveals that aesthetic attention bypasses this collapsed capacity entirely. One can attend to the quality of morning light not by first establishing that one’s life has meaning or that such attention serves some larger purpose, but simply because appearing “is open to us at all times,”¹⁷

9 Sheri Irvin, “The Pervasiveness of the Aesthetic in Ordinary Experience,” *British Journal of Aesthetics* 48, no. 1 (January 2008): 29.

10 Irvin, “The Pervasiveness of the Aesthetic,” 36.

11 Irvin, “The Pervasiveness of the Aesthetic,” 33.

12 Martin Seel, *Aesthetics of Appearing*, trans. John Farrell (Stanford University Press, 2005), 46.

13 Seel, *Aesthetics of Appearing*, 46.

14 Seel, *Aesthetics of Appearing*, 47.

15 Seel, *Aesthetics of Appearing*, 91.

16 Seel, *Aesthetics of Appearing*, 92.

17 Seel, *Aesthetics of Appearing*, 20.

demanding only presence to the here and now. In this phenomenological model of everyday aesthetics, Saito eliminates the requirement of achievement, Irvin eliminates the requirement of sustained attention, and Seel eliminates the requirement of justification itself.

Taken together, these three strands of aesthetic theory progressively dismantle barriers that would otherwise make engagement impossible during crisis. For someone who cannot trust any larger evaluative framework, these micro-engagements may be the only “reasons” that survive. Aside from premises in an argument for life, these reasons form concrete occasions to remain in touch with what is happening. They might be enough to get through the crisis moment, through the next hour, through the night, not because they prove anything about life’s ultimate worth but because they provide genuine reasons for engagement that persist independently of such proofs. Herein lies the practical import of recognizing aesthetic engagement outside evaluative structures, to continue without resolving the unresolvable and without that requiring capacities that crisis has compromised.

Suicidal crisis is not a passive state but an active one which must be continuously renewed through a repeated mental loop. The suicidal individual must constantly judge themselves deficient. Aesthetic attention breaks this loop not by introducing better reasoning, but by temporarily exiting the evaluative mode altogether. Instead of questioning the justification of their existence; an individual is simply perceiving. The question has not been answered; it has been abandoned. And in that absence, the immediate imperative of the crisis loses the cognitive fuel it requires to sustain itself. While this paper has focused on the recursive trap of ethical self-assessment as one prominent structure of suicidal crisis, the interruption that aesthetic attention provides may not be limited to this case. Whatever the specific content of the crisis logic, whether rooted in self-condemnation, unbearable pain, or perceived burden to others, the mechanism may operate similarly. Aesthetic attention does not engage the content of that logic but suspends the evaluative mode in which any such logic must operate.

Aesthetic attention, as this paper has argued, cannot be forced—crisis logic can sometimes be sufficiently dominant to foreclose the possibility of turning toward it entirely. Matt Haig, a writer who survived severe depression and suicidal crisis, documents both sides of aesthetic attention in relating to suicidal crisis in his memoir *Reasons to Stay Alive*. He describes standing at the edge of a cliff during a suicidal episode, looking out at what he calls the most beautiful view he had ever known and feeling nothing. The aesthetic experience was present, but his suicidal logic was too far gone to allow attention to turn toward it, occupied entirely with summoning the courage to jump. Yet the same author, from the other side of the crisis, offers his thoughts after he recovered: “Wherever you are, at any moment, try and find something beautiful. A face, a line out of a poem, the clouds out of a window, some graffiti, a wind farm. Beauty cleans the mind.”¹⁸ Haig’s trajectory suggests that aesthetic attention cannot be forced during the depths of crisis, but that its cultivation remains one of the most reliable forms of contact with existence available when evaluative reasoning has collapsed.

This proposal does not claim that aesthetic experience shows life to be meaningful, or that beauty reveals an underlying goodness of existence, or that recognizing such beauty entails a duty to go on. Those moves would simply reinstall the evaluative machinery that crisis has rendered inoperative. Instead, aesthetic engagement is offered as a way of continuing that does not depend on resolving whether continuing is justified. The person in crisis need not infer “because I noticed something beautiful, therefore my life has value”; the noticing itself is the act of remaining, moment by moment, in the world.

18 Matt Haig, *Reasons to Stay Alive*, (Penguin Books, 2015), 249.

In this sense, aesthetic engagement does not answer Camus’s question of whether life is worth living, nor does it refute those who conclude that it is not. It simply refuses the demand for an answer as a precondition of continuing to exist. As long as light, sound, and texture remain, so does the possibility of an engagement that is neither evaluative nor justificatory, but nevertheless real. For some individuals in suicidal crisis, that may be the only form of contact with existence available, and the only one that does not immediately collapse back into the very evaluative structures that have become unlivable.

Expanding Aesthetic Engagement with Existentialism

When Camus insists that “one must imagine Sisyphus happy,” he gestures toward a mode of endurance that is not reducible to moral duty or existential choice, but he leaves the nature of this happiness largely unspecified. His language of “revolt” and “lucidity” suggests an ongoing stance toward absurdity, yet he offers little account of the concrete practices that could sustain such a stance over time, especially for someone in crisis. The very image he chooses, of a man condemned to push a rock up a mountain eternally, contains, however, the seeds of a different answer than the one Camus explicitly develops.

If meaning cannot arise from hope for change or from retrospective narrative unity in Sisyphus’ case, then whatever sustains him must be located in the ongoing activity itself. Yet it seems insufficient to attribute this entirely to a conscious posture of absurdist rebellion; Camus gestures toward something more within the lived experience of the struggle, though he leaves its precise nature intriguingly underdeveloped. What remains is the immediate, embodied engagement with the task in the weight of the rock, the strain of muscles, the rhythm of breath, and the feel of the ground underfoot. This is precisely the register picked out by phenomenologically oriented everyday aesthetics.

Let us compare Sisyphus with a child kicking a rock down a sidewalk. While the child’s situation is temporary and leisurely compared to Sisyphus’ infinite labor, the dynamics of the activity are relatively the same: the contact of body and stone, the rock’s trajectory, the minor adjustments of force and angle, and the repetitive rhythm that emerges. That absorption exemplifies aesthetic attention to appearing in the how of the experience as it unfolds rather than the why of its justification. In such cases, the activity is complete, requiring no external purpose or evaluative endorsement.

On this reading, Sisyphus can be imagined as happy not primarily because he heroically affirms the absurd, but because he is capable of this kind of immediate, non-evaluative, sensory absorption in the task of pushing. The rebellion that Camus names but does not analyze would then consist of a mode of presence rather than a judgment about his condition. The rock’s resistance, the texture of stone under his hands, the warmth or chill of the air, and the cadence of effort and brief descent when the rock rolls back: all qualitative features of an experience to which Sisyphus can attend, rather than meanings he is tasked with creating. The labor remains meaningless in a global sense, but it is concretely present in a way that can sustain engagement without requiring that the entire situation be justified.

This suggests that existentialism, at least in Camus’s version, points toward but does not complete a turn from evaluative justification to concrete experience. Camus moves closer to lived experience, yet his emphasis on revolt and lucidity still leans on a demanding, ongoing stance that may outperform the resources available to someone in crisis. Aesthetic engagement completes this trajectory by eliminating the residual evaluative demand. Where the Stoic must judge that virtue has become impossible, the existentialist must take respon-

sibility for creating meaning, and even Camus's absurd hero must continually affirm revolt, aesthetic attention requires none of these. It asks neither that the person be virtuous, nor that they authentically choose, nor that they maintain a life-affirming stance toward absurdity. It asks only for minimal, local, sensorial presence: pushing the rock, feeling the rock, and noticing breath, light, weight, and texture as they appear.

The aesthetic framework does not oppose existentialism so much as extend one of its latent insights. Existentialism rightly turns philosophy's attention from abstract essence to lived existence; everyday aesthetics specifies how that existence can be inhabited when an individual is actively in crisis. It should be noted that the aesthetic framework proposed here is not intended as a complete philosophy of life, but as a resource for crisis specifically—a way of remaining in existence when evaluative approaches have failed, preserving the possibility of whatever justification may follow.

What remains when one cannot “choose meaning” or “rebel against the absurd” may be exactly what sustains Sisyphus in the image that has so captivated readers: an absorption in the aesthetic dimensions of bodily, temporal, and material experience that makes it possible to continue without first settling the question of whether continuing is justified.

Death-Oriented Aesthetics

The aesthetic framework defended here must be distinguished from aesthetic representations that make death itself appear compelling. Consider the Werther effect,¹⁹ named after Goethe's novel, which sparked imitative suicides across Europe. When suffering is aestheticized through compelling narrative, framed as tragic depth, authentic self-expression, or genuine feeling, suicide can become socially contagious. Certain aesthetic categories can make voluntary death appear formally attractive.

If aesthetic representations can make death compelling, what prevents the aesthetic framework defended here from doing the same thing in lived experience? I have argued for first-person aesthetic engagement with sensory phenomena as a response to suffering. But could this very framework make one's own death appear aesthetically necessary or compelling?

A paradigmatic example of this danger is explicit in writer Yukio Mishima's aesthetic theory and his 1970 ritual seppuku. For Mishima, genuine understanding of death required physical embodiment. In *Sun and Steel*, Mishima writes: “However much the closeted philosopher mulls over the idea of death, so long as he remains divorced from the physical courage that is a prerequisite for an awareness of it, he will remain unable even to begin to grasp it.”²⁰ Mishima believed that purely literary or intellectual engagement with mortality left him fundamentally incomplete as a person. He wrote extensively on death, but could never achieve the warrior's embodied understanding of it. Death appeared to him as the singular aesthetic and existential act in which his divided self might finally be brought into unity. Mishima's aesthetic ideal of harmony between art, body, and life does not just accompany his suicide as background context but progressively orients him toward it as the only act capable of fulfilling that ideal.

In 1970, Mishima committed suicide in the form of a choreographed coup attempt followed by ritual seppuku, timed to coincide with the completion of his final novel and

19 David P. Phillips, “The Influence of Suggestion on Suicide: Substantive and Theoretical Implications of the Werther Effect,” *American Sociological Review* 39, no. 3 (1974): 340–354.

20 Yukio Mishima, *Sun and Steel*, trans. John Bester (Kodansha International, 1970), 36. Originally published 1968.

performed when his body was at peak physical condition. Death became Mishima's ultimate aesthetic act. By committing seppuku in the manner of a samurai warrior, Mishima believed he could finally embody the ideals he had only written about. The suicide was designed to prove his genuine commitment to these values, or that he could live and die according to the samurai code rather than merely aestheticize it from a distance. We can call this orientation *death-oriented aesthetics*: death that appears compelling not in spite of aesthetic considerations but precisely because it seems to offer completion, coherence, and unity.

The objection is that if aesthetic categories can make such a death appear formally attractive, then aesthetic attention might be structurally dangerous for someone in suicidal crisis. If one can aestheticize the completeness of a finished life, why should that not be at least as compelling as attending to the textures of ongoing experience? On this view, aesthetics seems double edged. It can beautify continuation, but it can also beautify cessation by presenting death as the final brushstroke that allows the composition of a life to “make sense.”

The answer begins with temporal and evaluative structure. Death-oriented aesthetics privileges a retrospective standpoint that takes life as a whole and asks whether it achieves narrative or formal unity. It views existence as a work to be judged from outside and then recruits death as a permissible or even necessary means of securing coherence between values and actions, image and reality. In Mishima's case, his aesthetic demands to his writing and ethical demands to his bushidō code²¹ intertwined so tightly that only a spectacular, publicly visible death seemed adequate.

Aesthetic experience, like any deep human capacity, has a shadow side: the same sensitivity to beauty and form that can anchor someone to existence can, in other configurations, make death itself appear the more compelling aesthetic act. This paper has focused deliberately on the constructive dimension of aesthetic attention—its capacity to provide non-evaluative contact with existence—while recognizing that the darker possibility Mishima represents is real and cannot be fully theorized away. The everyday aesthetic framework developed earlier does not stand outside life to appraise it as a completed work. It attends instead to the appearing of phenomena in the continuous present: to light, sound, texture, and bodily sensation as they are given now. This mode of engagement is inherently incomplete and cannot be perfected, because its very form is ongoing responsiveness to what continues to show up. It does not ask whether a life is coherent or justified. It asks only for local, momentary contact with what is present.

This difference also helps address a further concern, namely that everyday aesthetic attention may be profound in crisis yet feel trivial or even banal once someone has recovered. That worry is understandable if aesthetic engagement is treated as a substitute for richer ethical, relational, or narrative goods. The earlier account does not claim that everyday aesthetics should become the highest value or the permanent center of a flourishing life. It presents aesthetic attention as a minimal, crisis-suitable mode of contact with existence that remains available when more demanding capacities for evaluation, planning, and self-narration have collapsed. In crisis, noticing small details of existence can be genuinely life preserving because it does not require that one first answer whether life is worth living. Outside crises, such noticing may recede into the background of richer projects, relations, and commitments without losing its significance as a basic way of inhabiting experience.

Everyday aesthetic engagement does not have to be life-changing: it can function as a

21 Hurst, “*Bushidō* Ideal.” A code of samurai ethics, or the “way of the warrior.” A samurai would commit seppuku (ritual embowelment) as an honorable way to die.

low-threshold practice that keeps open a channel of contact with the world when other forms of orientation have failed for suicidal individuals in crisis, and it can continue, in more ordinary times, as one among many ways of inhabiting a life that also includes ethical striving, narrative reflection, and robust agency. Death-oriented aesthetics becomes dangerous at the precise point where aesthetic categories are conscripted into the service of evaluation, justification, and narrative completion. It is that conscription that makes suicide appear as the only act capable of securing coherence, authenticity, or wholeness. The version of aesthetics defended here resists this conscription. It does not try to show that life as a whole is meaningful or justified, but offers instead a way of remaining in touch with what appears when the very demand for justification has become unlivable. In crisis, that may be the most that can be asked of any philosophical framework, and for some, it may be enough to allow continuation without first resolving whether continuing to live can be fully explained.

Works Cited

- Améry, Jean. *On Suicide: A Discourse on Voluntary Death*. Translated by John D. Barlow. Indiana University Press, 1999. Originally published in 1976.
- Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. Hackett, 1999.
- Camus, Albert. *The Myth of Sisyphus and Other Essays*. Translated by Justin O'Brien. Vintage Books, 1991. Originally published 1942.
- Dewey, John. *Art as Experience*. Perigee Books, 2005.
- Greenman, Craig M. "Cobain's Suicide Note as a 'Last Work': Art, Authenticity, and Ethics." In *Expression and Survival: Studies in the Aesthetics of Extreme Experience*, edited by Jonathan Mitchell, 145–166. Routledge, 2014.
- Haig, Matt. *Reasons to Stay Alive*. Penguin Books, 2015.
- Hurst, G. Cameron III. 1990. "Death, Honor, and Loyalty: The *Bushidō* Ideal." *Philosophy East and West* 40, no. 4 (October): 511–527. <https://www.jstor.org/stable/1399355>.
- Irvin, Sherri. 2008. "The Pervasiveness of the Aesthetic in Ordinary Experience." *British Journal of Aesthetics* 48, no. 1 (January): 29–44.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Edited and translated by Mary Gregor and Jens Timmermann. Cambridge University Press, 2012.
- Mishima, Yukio. *Sun and Steel*. Translated by John Bester. Kodansha International, 1970. Originally published in 1968.
- Nietzsche, Friedrich. "Twilight of the Idols." In *The Anti-Christ, Ecce Homo, Twilight of the Idols, and Other Writings*, edited by Aaron Ridley and Judith Norman, translated by Judith Norman. Cambridge University Press, 1997.
- Phillips, David P. "The Influence of Suggestion on Suicide: Substantive and Theoretical Implications of the Werther Effect." *American Sociological Review*, 39(3), 340–354, 1974.
- Saito, Yuriko. *Everyday Aesthetics*. Oxford University Press, 2007.
- Sartre, Jean-Paul. *Existentialism Is a Humanism*. Translated by Philip Mairet. Methuen, 1948.
- Sartre, Jean-Paul. *Being and Nothingness*. Translated by Hazel E. Barnes. Washington Square Press, 1993.
- Seel, Martin. *Aesthetics of Appearing*. Translated by John Farrell. Stanford University Press, 2005.
- Seneca. *Letters on Ethics: To Lucilius*. Translated by Margaret Graver and A.A. Long. University of Chicago Press, 2015.

Bare Life in the Technopoly: The Potential for Democracy Under Surveillance Capitalism in the Age of Artificial Intelligence

Rosalind Beattie

Abstract: Through a critical biopolitical and surveillance capitalist framework, artificial intelligence can be grounded as a technology manifesting within systems of power and influencing human life as an extension of these powers. I will develop the argument that these systems of power, which fundamentally reduce the individual to that of bare life, threaten the workings of a democratic society. By understanding the harms of artificial intelligence as it becomes increasingly integrated into societal structures, I propose a revival of democracy through an emphasis on the necessity of shared physical space outside of the digital realm.

Introduction

Artificial intelligence (AI) technologies are progressively being inserted into social systems, introduced to the public in marketable forms, and, most critically, embedded in our collective cultural consciousness. While these technologies appear to introduce a host of new ethical dilemmas, it is paramount to move forward with a conception of AI that is contextualized and historically grounded. While AI presents novelty in its complexity, technologies never materialize independently; rather, they always emerge in the context of the dominant power systems around them. The ethical issues present with AI technologies are not unique, but rather *new iterations* of the ways systems of power manifest in life. Combatting these issues must involve a revival of democracy through shared discourse in physical spaces rather than digital ones.

The Rise of the Technopoly

It is productive to understand AI as fundamentally a technology, even if its particular qualities make it novel. While not about the AI technologies of today, the work of American author and thinker Neil Postman provides a distinctly prescient understanding of technology

and culture, one that is particularly relevant to the current era’s grappling with AI technologies.

In his 1985 book *Amusing Ourselves to Death: Public Discourse in the Age of Showbusiness*, Postman writes in response to the influences of the entertainment-driven technologies of the time, namely radio and its extension into television. Postman tracks how the instantaneous nature of communication introduced by the telegram drastically molded the broader culture of the United States into one where “...all public discourse increasingly takes the form of entertainment.”¹ His emphasis on public discourse points to a prominent concern over the impact of technology on not merely culture in its broadest sense, but specifically the culture of democracy.

Central to *Amusing Ourselves to Death* is an understanding of conversation as “...all techniques and technologies that permit people of a particular culture to exchange messages,”² and that the connection of conversation to the Platonic notion that how we converse “...will have the strongest possible influence on what ideas we can conveniently express.”³ For Postman and for Plato, the *mode* of communication determines what kinds of messages and ideas are presented, and more importantly, the messages present in public discourse. The technologies of communication, by virtue of their form, will prioritize and promote certain messages over others. For Postman at the time of writing *Amusing Ourselves to Death*, this is the shortened, overly visual communication of television. For the average person in the current United States, it is the attention-grabbing, short-form content of most social media platforms. The role of data collection and its influence on these messages will become more relevant in the section exploring surveillance capitalism, but the importance of Postman’s message from 1985 is this—technology is designed to communicate information in a certain way, and this form of communication creates *the possible discourse* of a society.

The epistemological significance of this is not lost on Postman, who argues that “...a major new medium changes the structure of discourse; it does so by encouraging certain uses of the intellect, by favoring certain definitions of intelligence and wisdom, and by demanding a certain kind of content—in a phrase, by creating new forms of truth-telling.”⁴ Within Postman’s framework, the United States emerged initially as a country grounded in the written word, with the rise of television presenting not only a technological shift but an epistemological one. Truth that was once found in the written argument was then found in images and video, which, as extensions of sight, simultaneously presented an unchanging snapshot of time, while also opening the door to advertising and similar attempts at influencing consumers. With the rise of powerful generative AI capable of creating images and videos that are increasingly difficult to distinguish from the previous visual “truth” established by television, a new epistemological shift is underway. A photograph or video is no longer reliable evidence of something real and tangible. Thus, AI is currently obliterating the conception of truth based in visual communication—it is warping the way the world is known to the individual.

The effects of technological modes of communication on politics are not unique to the epistemological pollution of public discourse by the advent of television Postman describes;⁵

1 Neil Postman, *Amusing Ourselves to Death: Public Discourse in the Age of Showbusiness* (Methuen, 1985), 3.

2 Postman, *Amusing Ourselves to Death*, 6.

3 Postman, *Amusing Ourselves to Death*, 6.

4 Postman, *Amusing Ourselves to Death*, 27.

5 Postman, *Amusing Ourselves to Death*, 28.

social media, accompanied by its algorithms, has continued this rather disastrous muddying of the shared dialogue of democratic systems of governance. Michaelangelo Anastasiou explores the role of rhetoric in the age of political discourse that has largely moved to internet-based modes of communication. Anastasiou describes how, in relation to the often incoherent, sensationalized rhetoric of right-wing politics, “democratic forces are still ensnared in drawn-out analyses and discussions that appeal to the principles of ‘internal rationality’ and the intricacies of scientific observation.”⁶ Anastasiou illustrates how the leadership of the Democratic Party in the United States is stuck in bygone modes of communication like television and the written word, while the Republican Party has adapted to the instantaneous communication of the internet, and its messaging has shifted as such. The information given through a traditional press briefing is now accompanied by social media posts directly from politicians, with the latter paving the way for sensationalized disinformation. Bias has frequently accompanied the media, but the current unfiltered flood of access directly to political leaders poses a danger to the authority of expertise. While Anastasiou argues for the Democratic Party to lean into this rhetoric, it is perhaps worthwhile to take a step back and question *how the mode of communication creates the message*.⁷

Postman takes these baseline concepts from *Amusing Ourselves to Death* and develops them into a fleshed-out framework in his 1992 book *Technopoly: The Surrender of Culture to Technology*. For Postman, culture is being exceedingly driven by information and technological progress for the sake of progress, an issue that he says leads to information glut—a decontextualized, massive amount of information that can travel rapidly across the world.⁸ The breakneck pace at which these changes have occurred leaves culture and its institutions at a loss as to how to integrate such technologies back into itself. What is left is a culture defined by its technologies and the needless search for more information, which Postman coins as *technopoly*. For Postman, it was the lack of regulation over industry and the abandonment of traditionally held cultural beliefs that led to the United States becoming the first true technopoly.⁹

Even in 1992, Postman clearly describes the beginnings of a system of surveillance working itself into the private lives of people as a direct result of this shift into a technopoly. He asks how computer technology has “...been an advantage to the masses of people... Their private matters have been made more accessible to powerful institutions. They are more easily tracked and controlled; are subjected to more examinations; are increasingly mystified by the decisions made about them; they are often reduced to mere numerical objects.”¹⁰ The kinds of technological insights Postman presents here are alarmingly pertinent to the current digital age, where the production of artificial intelligence systems directly relies upon data collected from the individual. The fact of user surveillance has become so commonplace that it hums in the background of life, in the well-timed ads on a social media page, or the wordy terms and conditions presented on nearly every webpage and app. It is the *reduction of people into numerical objects* that will prove useful to explore from a biopolitical perspective and in conjunction with a larger capitalist structure.

6 Michaelangelo Anastasiou, “The Hegemonic World Picture: Representation, Post-Truth, and Artificial Intelligence,” *Philosophy & Social Criticism* 0, no. 0 (2025): 16, <https://doi.org/10.1177/01914537251374951>.

7 Postman, *Amusing Ourselves to Death*, 10.

8 Neil Postman, *Technopoly: The Surrender of Culture to Technology* (Vintage Books, 1992), 61.

9 Postman, *Technopoly*, 53.

10 Postman, *Technopoly*, 10.

Surveillance Capitalism and Control

While Postman hints at the role of capitalism in the development of technology, Shoshana Zuboff presents a framework for understanding the new ways that capitalism manifests in the digital age, coining the term *surveillance capitalism*. Zuboff outlines this surveillance capital framework as it was initially introduced by Google and quickly adopted by various digital corporate enterprises.¹¹ Google was able to sell the surplus data generated from search queries, creating a market that "...was not an exchange with users but rather with companies that understood how to make money from bets on users' future behavior...users were no longer ends-in-themselves...Instead, users are the *human natural* source of free raw material that feeds a new kind of manufacturing process designed to fabricate *prediction products*."¹² AI fits neatly into this market as it fundamentally relies on this human-generated data, and big tech companies utilize AI systems to more accurately predict the actions of the users they surveil. AI being introduced to the public in the form of chatbots, search engines, and the like, is merely the next iteration of surveillance capitalists attempting to maximize profit. Just as Google took user data and turned it into a raw product to sell to advertisers, this same raw data is now being repackaged into a new generation of products to sell back to the user, all without compensation, voice, or autonomy to the people who are exploited in the process.

The surveillance capital economic model has a certain level of opacity necessary to its function. As Zuboff writes, "economies of action go further. These new systems and procedures take direct aim at individual autonomy, systematically replacing self-determined action with a range of hidden operations designed to shape behavior at the source."¹³ What is even more profitable than highly targeted advertising is the ability to actively *control* the future of consumer behavior. Google won't just sell a *susceptible* customer to advertisers—they can sell a *guaranteed* customer. In *An Inquirer's Guide to Ethics in AI*, authors Matthew Silk and Ian MacDonald describe "nudging," where a decision is presented in such a way that it psychologically favors one choice over another,¹⁴ and subsequent "hypernudging," which "...uses AI to continuously reconfigure the nudge based on data that an algorithm is getting from you and others."¹⁵ Hypernudging with highly complex algorithms has the potential to alter what people see online in real time. This is all done with little to no regard for the autonomy of the users exploited in this system. Users are essentially manipulated, which undermines the democratic system that, on the most fundamental level, relies on the autonomous ability of the populace to think for themselves. When this ability is compromised in such a subtle fashion, it undermines a society driven by rational choice and deliberation.

Zuboff describes how "...populations are conceptualized as undifferentiated 'users,' who are merely the sources of raw material for a digital-age production process aimed at a new business customer...social relations are no longer founded on mutual exchange."¹⁶ The notion of mutual exchange being necessary for social relations is a compelling idea in relation to democratic governance. Élodie Marchand, Théo Durand, and Maëlle Dubois tackle this is-

11 Shoshana Zuboff, "Surveillance Capitalism and the Challenge of Collective Action," *New Labor Forum* 28, no. 1 (2019): 14, <https://doi.org/10.1177/1095796018819461>.

12 Zuboff, "Surveillance Capitalism," 13.

13 Zuboff, "Surveillance Capitalism," 18.

14 Matthew S.W. Silk and Ian J. MacDonald, *An Inquirer's Guide to Ethics in AI* (Broadview Press, 2024), 198.

15 Silk and MacDonald, *An Inquirer's Guide*, 199.

16 Zuboff, "Surveillance Capitalism," 21.

sue, describing how, from a marketing perspective, the highly individualized feedback loops of AI-driven advertising "...could lead to a world where public discourse even around brands or products is splintered—fewer unifying cultural moments, more niche microcultures."¹⁷ Even if one is to accept capitalist markets as permissible, embedding AI into structures like marketing leads to a narrowing of the world and a loss of connectivity with others. Completely personalized advertisements drastically limit the choices a consumer can make and present different versions of entertainment and products to each person on an individual level. This results in a loss of word-of-mouth recommendations and shared discourse about brands and products unmediated by technology. Marketing often attempts subtlety, but there is a striking difference between colorful packaging in the supermarket and an algorithm that can predict the most opportune moment when the user is most susceptible to an advertisement.

It is worth noting that Zuboff would perhaps see the root of the issues of AI technologies as lying solely with the economic power systems around them, rather than with something inherent to the technologies themselves. She writes, "The primacy of economics over technology is not new, but capitalism has long found it useful to confound society by concealing itself within the Trojan horse of technology, in order that its excesses might be perceived as the inexorable expression of the machines it employs."¹⁸ While it is true that technology is something that can be weaponized by those in power, Postman provides a more nuanced understanding where technologies are not neutral, but visions for possibilities of the future. For Postman, technology and its integration into daily life change the way certain values are understood by shifting the context around them, such as the advent of television leading to a valuing of entertainment over reliable information.¹⁹ The way that a technology is designed is essential to how it will impact society.

This does not mean that Postman is dismissive of the role of capitalism in technological development. If anything, he would argue that the digital milieu of the current day, and all the technologies it employs, were developed with capitalism in mind. In his 1998 talk "Five Things We Need to Know About Technological Change," Postman states that "The consequences of technological change are always vast, often unpredictable, and largely irreversible. That is also why we must be suspicious of capitalists. Capitalists are by definition not only personal risk takers but, more to the point, cultural risk takers."²⁰ Technology is rarely a factor that is added or subtracted while avoiding unforeseen, far-reaching consequences.²¹ Those who invent technologies are the arbiters in shaping the *values* of society, yet they often do so impulsively and with little regard for the permanency of these changes. The introduction of smartphone technology created unprecedented user access to the world, while at the same time halting in-person interactions. It is not obstinate to recognize that in public settings, the majority of people focus their attention on their phones rather than on those around them. This is a change that is unlikely to go away, given that a smartphone is nearly necessary to effectively navigate society. Suffice to say, the one driving force for the capitalist is the

17 Marchand, Élodie, Théo Durand, and Maëlle Dubois. "Surveillance Capitalism, Consumer Subjectivity and Marketing in the Age of Artificial Intelligence." *Social Science Chronicle* 5, no. 1 (2025): 8, <https://doi.org/10.56106/ssc.2025.001>.

18 Zuboff, "Surveillance Capitalism," 12.

19 Postman, *Amusing Ourselves to Death*, 7.

20 Neil Postman, "Five Things We Need to Know about Technological Change" (lecture, Denver, CO, March 28, 1998).

21 Postman, *Technopoly*, 5.

accumulation of capital, and this easily lends itself to the collection of power and control over the market. It is all too easy for the capitalist to integrate the individual into the market *as an object* if doing so presents a compelling economic interest. It is when the capitalist is left unregulated, with complete control over the market, that this *reduction of human life* takes on a biopolitical significance.

Biopower and the Apparatus

Zuboff astutely describes how the result of this surveillance capitalist system “...has been an involuntary merger of personal necessity and economic extraction, as the same channels that we rely on for daily logistics, social interaction, work, education, health care, access to products and services, and much more, now double as supply chain operations for surveillance capitalism’s surplus flows.”²² It is becoming increasingly difficult to function without relying on technologies of some kind that are integrated into a surveillance capital system, whether it be through a bus pass that is only available on an app, an automated customer service agent answering the phone, or an AI summary embedded in a previously used technology.

The technologies of surveillance capitalism, and by extension AI, have become unavoidable in daily life. Not only this, but the powers that implement these technologies have a distinct interest in the private lives of individuals. As Zuboff states, “surveillance capital wants more than your body’s coordinates in time and space. Now it violates the inner sanctum, as machines and their algorithms decide the meaning of your sighs, blinks, and utterances; the pattern of your breathing and the movements of your eyes; the clench of your jaw muscles; the hitch in your voice.”²³ While Zuboff’s description appears exaggerated, one has to look no further than the use of Flock cameras²⁴ and similar surveillance technologies to grasp the extent of surveillance capitalism’s invasion into the private sphere. Regarding camera surveillance, Jay Stanley, policy analyst for the ACLU, describes the development of “...algorithms that can detect people or behaviors that are ‘unusual,’ ‘abnormal,’ ‘deviant,’ or ‘atypical.’ Finding statistical deviations may be an interesting mathematical challenge, but when that shades into finding deviant people it should raise alarms.”²⁵ From determining a customer’s buying patterns based on grocery store footage to policing the movements of a crowd of protesters, the companies developing surveillance technologies stand to profit from dissecting the minutiae of human behavior. The technologies of the surveillance capitalists become embedded in social life to the extent that their power feels indistinguishable from the power of formal governance, often because the government invests in and implements these very surveillance technologies. This power exerts itself blatantly and completely on the individual and how they live.

Giorgio Agamben extrapolates on Michel Foucault’s conception of *biopower*, which can be understood as “...what brought life and its mechanisms into the realm of explicit calculations and made knowledge-power an agent of transformation of human life.”²⁶ Biopower

is essentially the various methods that the State employs to control large populations, often with a focus on how life is lived. In this definition, the connection is already apparent between biopower and the mechanisms by which surveillance capitalism functions. However, Agamben develops this understanding further. For Agamben, “...the development and triumph of capitalism would not have been possible...without the disciplinary control achieved by the new bio-power, which, through a series of appropriate techniques, so to speak created the ‘docile bodies’ that it needed.”²⁷ In order for human labor to be extracted and exploited under a capitalist structure, the individual must be seen as a cog in society whose life can be regulated, controlled, and disciplined by the State. It is apparent how the reduction of human life into mere data for surveillance capitalism is an iteration of the biopower at play that capitalism more generally relies on, a creation of docile bodies in the form of the user who generates the data surveillance capitalism requires.

For Foucault and for Agamben, biopower is not merely a representation of the State’s power. Biopower goes beyond the classical models of sovereignty and instead embeds itself into much larger frameworks. In part, this involves “...*technologies of the self* by which processes of subjectivization bring the individual to bind himself to his own identity and consciousness and, at the same time, to an external power.”²⁸ There is an element of self-monitoring to biopower, one that is far more subtle than the explicit, forceful power from the traditional State. Biopower plays out in social norms, internalized by the individual. It is in the way that young girls learn beauty standards, that gender is performed through clothing or mannerisms, and the expectation to maintain the health of the body to maximize productivity.

Author Mikkel Keni Bruun explores the use of wearable devices as modes of self-surveillance. As Bruun describes it, “the surveilling gaze has, in an important sense, been inverted—turned inwards—as we are increasingly encouraged and required to watch our selves.”²⁹ These are the exact kinds of technologies of the self that Foucault and Agamben describe; it is the kind of power that manifests itself as a self-imposed, inward surveillance. In the case of wearable devices such as Fitbits, Apple Watches, or Oura rings, companies present a self-surveillance technology justified as a tool to optimize mental and physical health, and the body becomes an object that requires monitoring to meet normative standards of health.³⁰ All the while, the highly personal data that is collected by these technologies is at best kept in a data center that could potentially be accessed by threat actors, and at worst, is free to use by these companies for their own capital gain. While user agreements are implemented by these companies to avoid potential legal repercussions, such agreements are often lengthy and filled with language that is difficult for the user to understand. Thus, such agreements rarely feel like a fair interaction or exchange between the user and the company.

It is this form of biopower operating through technologies of the self that Agamben relies heavily on in his larger framework, where he establishes “bare life” as that which is included in the social order only through its exclusion.³¹ As Agamben writes, “in such a space of exception, subjection to experimentation can...either return the human body to life...or

Pantheon Books, 1978), 143.

27 Giorgio Agamben, *Homo Sacer: Sovereign Power and Bare Life* (Stanford University Press, 1998), 3.

28 Agamben, *Homo Sacer*, 5.

29 Mikkel Keni Bruun, “Watching Our Selves: Fitness and Mindfulness as Practices of Self-Monitoring in Britain,” in *Rhythm and Vigilance* (Policy Press, 2025), 26.

30 Bruun, “Watching Our Selves,” 27.

31 Agamben, *Homo Sacer*, 29.

22 Zuboff, “Surveillance Capitalism,” 25.

23 Zuboff, “Surveillance Capitalism,” 17.

24 Jay Stanley, “Flock’s Aggressive Expansions Go Far Beyond Simple Driver Surveillance,” *Free Future* (blog, American Civil Liberties Union, August 18, 2025), <https://www.aclu.org/news/privacy-technology/flock-roundup>.

25 Jay Stanley, *The Dawn of Robot Surveillance: AI, Video Analytics, and Privacy* (American Civil Liberties Union, June 13, 2019), <https://www.aclu.org/publications/dawn-robot-surveillance>.

26 Michel Foucault, *The History of Sexuality Volume 1: An Introduction*, trans. Robert Hurley (Reprint,

definitively consign it to the death to which it already belongs. What concerns us most of all here, however, is that the physician and the scientist move in the no-man's-land into which at one point the sovereign alone could penetrate."³² Agamben's concern is not merely over the consequences of the traditional sovereign holding power over life itself. It is when this sovereignty is expressed in non-governmental entities, when biopower is delivered in all the nuances of the bureaucratic structures, that the individual is reduced so completely to bare life, to simply the biological state of being alive rather than a qualified, political existence. Instead of the physicians or scientists, it is the surveillance capitalists who have now entered this no-man's-land of the sovereign. Users are taken merely as living beings, whose actions can be manipulated in order to extract further capital.

Agamben sheds light on how the State became so invested in the life of the individual through the use of the term "apparatus." Agamben modernizes Foucault's earlier conception of the apparatus, understanding it as "...anything that has in some way the capacity to capture, orient, determine, intercept, model, control, or secure the gestures, behaviors, opinions, or discourses of living beings."³³ In this light, the modes of communication in the technopoly, including the surveillance technologies employed by capitalists, are all forms of apparatuses that influence and shape life. Agamben even states that it is possible to define "...the extreme phase of capitalist development in which we live as a massive accumulation and proliferation of apparatuses."³⁴ Viewing AI technologies through the lens of the apparatus is particularly helpful for understanding how technologies come to impose themselves on life in a surveillance capitalist system.

What the current biopolitical system of surveillance does is, fundamentally, to justify total control with the most contradictory claims to the reality of the situation. Agamben writes that "it is only an apparent paradox that the harmless citizen of postindustrial democracies... who readily does everything that he is asked to do, inasmuch as he leaves his everyday gestures and his health, his amusements and his occupations, his diet and his desires, to be commanded and controlled in the smallest details by apparatuses, is also considered by power... as a potential terrorist."³⁵ The individual caught up in the technological apparatuses of our time is a normal, predictable human to be collected in datasets and calculated; yet, they are somehow equally a potential terrorist and threat to the State, which justifies such surveillance.

This kind of surveillance can be classified as a *banopticon*, where the state of emergency invoked by the government involves the use of surveillance technology to categorize individuals and evaluate their potential threat.³⁶ As Guilherme Giantini describes, "...the use of personal databases did not remain exclusive to...the intent of maintaining social order. In the era of mass communication media, political instability, and consumerism, personal data started to draw attention to occurrences in which their instrumentalization by powerful AI processing tools started attending interests from political, economic, religious, and gov-

32 Agamben, *Homo Sacer*, 159.

33 Giorgio Agamben, *What Is an Apparatus?* (Stanford University Press, 2009), 14.

34 Agamben, *What is an Apparatus?*, 15.

35 Agamben, *What is an Apparatus?*, 23.

36 Guilherme Giantini, "The Sophistry of the Neutral Tool. Weaponizing Artificial Intelligence and Big Data into Threats toward Social Exclusion," *AI and Ethics* 3(2023): 1052, <https://doi.org/10.1007/s43681-023-00311-7>.

ernmental spheres..."³⁷ Technologies that were previously implemented under the guise of public safety and social order have since developed into a corporate market. If surveillance capitalism has become so embedded in the life of the individual, it is a result of the largely unregulated technologies that were allowed to saturate the private lives of people.

AI, Power, and Democracy

Agamben emphasizes power's infringement on private life because of its profound implications for political life. For Agamben, "every attempt to rethink the political space of the West must begin with the clear awareness that we no longer know anything of the classical distinction between *zoe* and *bios*, between private life and political existence, between man as a simple living being at home in the house and man's political existence in the city."³⁸ Life itself becomes an issue of the State, and in an unregulated surveillance capitalist system, an issue of large tech companies that conduct this surveillance. The individual citizen is reduced to bare life, to being a docile body that generates data which can be extracted for corporate profit. In its current form, Artificial intelligence is not and can never be a beneficial, neutral tool for a democratically aligned society if only it were "used correctly" because it is functioning as intended; AI is designed to see people as mere objects among objects, without any of the individual autonomy necessary for productive discourse.

Many thinkers have been taking up the mantle of envisioning a democratic way forward under these systems of power and control. Ashish Makanadar expands on surveillance capitalism as a framework, proposing a reclaiming of data sovereignty by "empowering everyday citizens with user-friendly tools for encrypting communications, obfuscating data trails, and revoking data ownership claims by corporations,"³⁹ alongside legal reform from scholars "... building legal regimes for algorithmic due process and machine rights."⁴⁰ Similarly, Jingduo Hu advocates for communist digital communities with "...online spaces organized and governed collectively with shared ownership and participatory decision making rather than profit driven imperatives."⁴¹ Joseph Jones emphasizes the need for a care-based democracy, one that is "...concerned with everyday life, not just the official institutions of power."⁴² Life itself must be reclaimed as something valuable rather than merely a means to a profitable end if a democratic future is to be a possibility.

Sylvie Delacroix suggests the use of AI to create conversational spaces for democracy, a claim that could be potentially questioned if a complete refusal of AI as it stands is necessary in moving forward. As has been discussed, artificial intelligence is not a neutral tool that is being used for the wrong ends—it is developed with a certain view of the world in mind. As it stands, it is not possible to utilize current AI technologies in a "correct" way that will benefit democratic forms of governance and a sense of civic virtue. This is precisely because AI

37 Giantini, "The Sophistry of the Neutral Tool," 1052.

38 Agamben, *Homo Sacer*, 187.

39 Ashish Makanadar, "Digital Surveillance Capitalism and Cities: Data, Democracy and Activism," *Humanities and Social Sciences Communications* 11, no. 1 (2024): 2, <https://doi.org/10.1057/s41599-024-03941-2>.

40 Makanadar, "Digital Surveillance Capitalism and Cities," 3.

41 Jingduo Hu, "The Digitization of Life: The Biopolitics in the Era of Digital Capitalism." *World Futures* 81, no. 7 (2025), <https://doi.org/10.1080/02604027.2025.2563492>.

42 Joseph Jones, "Don't Fear Artificial Intelligence, Question the Business Model: How Surveillance Capitalists Use Media to Invade Privacy, Disrupt Moral Autonomy, and Harm Democracy," *Journal of Communication Inquiry* 49, no. 1 (2024): 20, <https://doi.org/10.1177/01968599241235209>.

technologies are fundamentally immersed and grown against the backdrop of the biopolitical and capitalist power structures that define them. However, as Delacroix notes, “democratic contexts require technological systems that can sustain hermeneutic uncertainty as a productive condition rather than treating all uncertainty as a technical deficit to be eliminated.”⁴³ The current technologies in use, our current apparatuses, are not designed democratically, so we must invent structures that are. Technology that is designed to manipulate and shape human behavior, to impair human decision-making, fundamentally goes against a democracy that values informed choices and collective deliberation. At the very least, people should be clearly informed about how their data is used and have alternative choices in technology without it impeding their ability to function in society.

Democracy is not something that can be flippantly left in the hands of capitalists. One need only look towards these new technologies and the kinds of goals and purposes they serve. As Shannon Vallor describes, “today’s AI systems trained on human thought and behavior are not minds. They are their own new kind of thing—something like a mirror. They don’t produce thoughts or feelings any more than mirrors produce bodies. What they produce is a new kind of reflection.”⁴⁴ Vallor’s mirror metaphor is productive in grounding artificial intelligence as a distinctly *human* technology. It is not some transcendent tool that will reveal a completely objective reality of the world, or some kind of universal *truth*. Instead, AI technologies are a series of highly complex algorithms used in a multitude of apparatuses that are built on the production of raw data *by humans*.

Mark Ryan evaluates three separate accounts of trust and determines that AI only meets the rational account, which is essentially a form of reliance.⁴⁵ The rational account of trust involves trust based on rationality, where “trust is simply a matter of prediction by the trustor, rather than being concerned about the trustee’s motivation.”⁴⁶ AI is similar, though more problematic, to something like a car; it is a technology that can be used and be potentially helpful for certain tasks, but in which no legitimate trust can be placed, because true trust involves the motivations of the trustee. As Ryan writes, “referring to AI as trustworthy would inappropriately elevate AI, while disavowing the responsibility of those developing and implementing it...Companies should be held responsible for impacts of their AI and should instil measures to avert harmful impacts...Responsibility should also lie with governments, industry leaders, research institutions and AI practitioners.”⁴⁷ By critically grounding artificial intelligence as a tool created by powerful actors and with a certain *view of the future* in mind, responsibility should correctly be placed on the shoulders of those who implement it in public life.

The fragility of the surveillance capitalist model lies in its reliance on isolation, which enables it to create individualized, distorted conceptions of reality. The user must be convinced that their digital environment is preferable to the vibrancy of the tangible world around them and the people in it. While Zuboff notes that “there can be no exit from processes that

43 Sylvie Delacroix, “Designing with Uncertainty: LLM Interfaces as Transitional Spaces for Democratic Revival,” *Minds and Machines* 35, no. 4 (2025): 19, <https://doi.org/10.1007/s11023-025-09736-x>.

44 Shannon Vallor, *The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking* (Oxford University Press, 2024), 39.

45 Mark Ryan, “In AI We Trust: Ethics, Artificial Intelligence, and Reliability,” *Science and Engineering Ethics* 26, no. 5 (2020): 6, <https://doi.org/10.1007/s11948-020-00228-y>.

46 Ryan, “In AI We Trust,” 4.

47 Ryan, “In AI We Trust,” 14.

are intentionally designed to bypass individual awareness and on which we must depend for effective daily life,”⁴⁸ it is also the case that a great deal of effort and manipulation must go into convincing the populace of the necessity of these processes. By revitalizing democracy and placing the burden of responsibility back on the government and the companies implementing AI technologies, there is the potential to overcome the surveillance capitalist model that life has become so entrenched in.

Democracy amounts to a committed way of interacting with people and a dedication to forming a society as such. According to John Dewey:

A democracy is more than a form of government; it is primarily a mode of associated living, of conjoint communicated experience. The extension in space of the number of individuals who participate in an interest so that each has to refer his own action to that of others, and to consider the action of others to give point and direction to his own, is equivalent to the breaking down of those barriers of class, race, and national territory which kept men from perceiving the full import of their activity.⁴⁹

Democracy necessitates direct communication and association with others in a way that is directly antithetical to the resigned mode of existence demanded by surveillance capitalism. Artificial intelligence can only distort reality and relegate the individual to bare life so long as the physical spaces of democratic discourse are abandoned. If technology is to be a tool of democracy, it must originate from a physical space of discourse and deliberation.

If, as Postman says, “...embedded in every tool is an ideological bias, a predisposition to construct the world as one thing rather than another, to value one thing over another, to amplify one sense or skill or attitude more loudly than another,”⁵⁰ then the technologies integrated into collective society must be reshaped through discourse in intentionally *real-world* spaces. Through reviving physical spaces of democracy where individuals interact face-to-face with those who are different from them, there is the possibility of creating new technologies that have a decidedly different constructed world in mind than the ones the populace is currently relegated to. The tools we use hold power, but so too does the decision-making that goes behind those tools. By beginning squarely in a democratically aligned physical realm of debate, new technologies can be envisioned in place of the old, and surveillance capitalism can be undone.

Works Cited

Agamben, Giorgio. *Homo Sacer: Sovereign Power and Bare Life*. Stanford University Press, 1998.

———. *What Is an Apparatus?* Stanford University Press, 2009.

Anastasiou, Michaelangelo. “The Hegemonic World Picture: Representation, Post-Truth, and Artificial Intelligence.” *Philosophy & Social Criticism* 0, no. 0 (2025): 1-23 <https://doi.org/10.1177/01914537251374951>.

Delacroix, Sylvie. “Designing with Uncertainty: LLM Interfaces as Transitional Spaces for Democratic Revival.” *Minds and Machines* 35, no. 4 (2025). <https://doi.org/10.1007/s11023-025-09736-x>.

48 Zuboff, “Surveillance Capitalism,” 25.

49 John Dewey, *Democracy and Education* (New York: Macmillan, 1916), 67.

50 Postman, *Technopoly*, 13.

- Dewey, John. *Democracy and Education*. Macmillan, 1916.
- Giantini, Guilherme. “The Sophistry of the Neutral Tool: Weaponizing Artificial Intelligence and Big Data into Threats toward Social Exclusion.” *AI and Ethics* 3 (2023): 1049–61. <https://doi.org/10.1007/s43681-023-00311-7>.
- Hu, Jingduo. “The Digitization of Life: The Biopolitics in the Era of Digital Capitalism.” *World Futures* 81, no. 7 (2025): 508–26. <https://doi.org/10.1080/02604027.2025.2563492>.
- Jones, Joseph. “Don’t Fear Artificial Intelligence, Question the Business Model: How Surveillance Capitalists Use Media to Invade Privacy, Disrupt Moral Autonomy, and Harm Democracy.” *Journal of Communication Inquiry* 49, no. 1 (2024). <https://doi.org/10.1177/01968599241235209>.
- Makanadar, Ashish. “Digital Surveillance Capitalism and Cities: Data, Democracy and Activism.” *Humanities and Social Sciences Communications* 11, no. 1 (2024). <https://doi.org/10.1057/s41599-024-03941-2>.
- Marchand, Élodie, Théo Durand, and Maëlle Dubois. “Surveillance Capitalism, Consumer Subjectivity and Marketing in the Age of Artificial Intelligence.” *Social Science Chronicle* 5, no. 1 (2025): 01-18. <https://doi.org/10.56106/ssc.2025.001>.
- Peacock, Vita, Mikkel Kenni Bruun, Claire Elisabeth Dungey, and Matan Shapiro. *Rhythm and Vigilance*. Policy Press, 2025.
- Postman, Neil. *Amusing Ourselves to Death: Public Discourse in the Age of Showbusiness*. Methuen, 1985.
- . *Technopoly: The Surrender of Culture to Technology*. Vintage Books, 1992.
- . “Five Things We Need to Know about Technological Change.” Lecture, Denver, CO, March 28, 1998.
- Ryan, Mark. “In AI We Trust: Ethics, Artificial Intelligence, and Reliability.” *Science and Engineering Ethics* 26, no. 5 (2020): 6. <https://doi.org/10.1007/s11948-020-00228-y>.
- Silk, Matthew S.W., and Ian J. MacDonald. *An Inquirer’s Guide to Ethics in AI*. Broadview Press, 2024.
- Stanley, Jay. “The Dawn of Robot Surveillance: AI, Video Analytics, and Privacy.” American Civil Liberties Union, June 13, 2019. <https://www.aclu.org/publications/dawn-robot-surveillance>.
- . “Flock’s Aggressive Expansions Go Far Beyond Simple Driver Surveillance.” *Free Future* (blog), August 18, 2025. <https://www.aclu.org/news/privacy-technology/flock-roundup>.
- Vallor, Shannon. *The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking*. Oxford University Press, 2024.
- Zuboff, Shoshana. “Surveillance Capitalism and the Challenge of Collective Action.” *New Labor Forum* 28, no. 1 (2019): 10–29. <https://doi.org/10.1177/1095796018819461>.

