Editor-in-Chief: Ziyue Li
Academic Editor: Weizhong Li
Art Editor: Yiyao Feng

NOUS REVIEW

A student-founded humanities review for young writers thinking beyond the classroom

Inaugural Issue October, 2026



Cinema as the Third Interpreter
by Ziyue Li

How Conversations Change Us
by Weizhong Li

The Bamboo Stove and the Poetry Society
by Yiyao Feng

The Attribution Default After LLMs
by Letian Cheng

Transpacific Sensory Politics
by Haoyan Shi

Semantic Stability and Floating Meaning: Thinking about the Language of AI
by Shenzhi Zhang

NOUS REVIEW

A JOURNAL OF IDEAS

A Letter from the Editors

Dear Readers,

There is a particular kind of hope that comes with beginning something new. When we first envisioned NOUS Review, we did not set out to build a credential or a competition. We wanted to create a place where young people like us could take ideas seriously and be taken seriously in return. The conversations that sustain the humanities are patient, open and honest, carried out with rigor, compassion and curiosity. We believed that secondary-school students could join these conversations not as mere spectators but as participants. This inaugural issue is our attempt to show that they can, and to invite many more of them to try.

The essays gathered here were written in a climate that does not make such conversations easy. Artificial intelligence now writes fluently on almost any topic. Several of our authors took this as an occasion to ask what it means for a mind to understand, how language derives meaning, and how far our recognition of other minds should extend. Meanwhile, social media encourages us to present a finished, curated version of ourselves; pushing against this narrative, other essays consider how the self is actually formed through deep conversations, gender, and foundational bonds of a family. Political-wise, life has hardened into opposing camps, where a resurgent conservatism and nationalism have made many debates feel settled before they even begin. In response to these polarizations, several authors turned to the conditions of public life, from the duty to tell the truth to the difficulty of changing entrenched institutions, while others trace how art, film and literary culture, from Ming China to twentieth-century America, have decided who is heard and who remains unseen.

That these questions are being raised by high school students matters. Theirs is the first generation to grow up entirely inside social media and to come of age alongside generative AI, and it is often described as the passive subject of both. These essays suggest otherwise. Their authors prove themselves more capable of examining the forces that shape them, of holding competing views in tension, and of treating disagreement as the beginning of inquiry rather than the end of it.

This is what our name is meant to express. In Greek philosophy, *nous* denotes the intellect: the capacity by which the mind grasps what is true and intelligible. Anaxagoras made *nous* the principle that brings order to the cosmos. For Plato and Aristotle, it was the highest activity of the soul, the faculty that sees first principles and understands rather than merely knows. We chose the name not because we claim to possess *nous*, but because it describes what we hope to cultivate. Today, in an age where information is abundant yet judgment is scarce, the capacity to understand, and not merely to react, is worth defending. We want NOUS Review to be a place where young scholars practice that capacity together, across schools, countries and traditions, and carry it into the public life they will soon shape.

Finally, we thank the authors in this issue for the seriousness and courage of their work. We are deeply grateful to our referees, whose careful reading and generous feedback gave these essays the attention they deserved. We also thank every member of our editorial team, who built this journal from nothing. We offer this first issue as the opening salvo of an intellectual dialogue marked by sincerity and humility, an inquiry into questions that no single answer completes, and a new chapter in a tradition of inquiry far older than ourselves, one we are humbled to inherit and hopeful to pass on.

Sincerely,

The Editors, NOUS Review

Contents

A Letter from the Editors.....i

Review & Critique

Should We Treat Nonhuman Animals Well Because They Have Rights, Interests, Neither, or Both?.. 1
Shenzhi Zhang

Two Junior Reviews on Beauvoir and Woolf..... 14
Xichen Liu and Wency Xu

A Home at the End of the World (2004): It's Just Love, Man. It's Nothing to Fear..... 25
Olivia Liu

On the Ending Scene of *Pygmalion* and *My Fair Lady*.....31
Fanxiao Sun

Lies Matter—The Punishment of Lying Politicians.....37
Yiming Chen

Yao Ming Resigned and His CBA Reform: Destiny or Inadequacy?..... 47
Qunyan Wang

Beyond Russell's Formula: A Critical Reassessment of Rousseau and Locke's Influence on Hitler, Roosevelt, and Churchill.....55
Minghao Liu

Research Papers

Semantic Stability and Floating Meaning: Thinking About the Language of AI..... 64
Shenzhi Zhang

The Attribution Default After LLMs.....78
Letian Cheng

How Conversations Change Us: Supplying a Hermeneutic Mechanism for Dover's Conversational Self..... 93
Weizhong Li

Cinema as the Third Interpreter: Interpretation, Recognition, and Misrecognition in *Gaslight*, *The Philadelphia Story*, and *Little Women*..... 113
Ziyue Li

Transpacific Sensory Politics: The Translation of Chinese Revolutionary Symbols and the Resistance Aesthetic of African American Artists (1930s–1970s).....129
Haoyan Shi

The Bamboo Stove and the Poetry Society: Social Networks and Political Contexts Behind the Literary Activities of the Wuxi Gentry in the Mid-Ming Period..... 154
Yiyao Feng

Feminism in Translation: Sociopolitical Context, Feminist Cinema, and Chinese Reception of *Barbie, Portrait of a Lady on Fire, and Night and Fog*..... 176
Ziyue Li

Review & Critique

Should We Treat Nonhuman Animals Well Because They Have Rights, Interests, Neither, or Both?

Shenzhi Zhang
Grade 11

Editors' Introduction

This essay asks why nonhuman animals deserve moral consideration, at a time when institutions increasingly recognize animal sentience while still organizing animal life around human purposes. It weighs interest-based ethics, associated with Bentham and Singer, against Tom Regan's rights-based view, and argues for a hybrid account in which each corrects the weaknesses of the other. The essay then makes a second, more original move. Drawing creatively on the social and political thought of Foucault and Agamben, it shows how the structures that manage and define animals can obscure their value, and it proposes the cultivation of moral attention as a necessary complement to the recognition of rights and interests. In doing so, it addresses problems often overlooked in discussions of animal ethics.

The essay pursues two related arguments, one concerning the grounds of animals' moral standing and the other concerning the structures that obscure it. The relation between its individual-level account of moral attention and its institutional critique is suggested rather than fully worked out, and its use of Foucault and Agamben is concise, gesturing toward the analogy between their cases and the case of animals without exploring it at length.

1. Introduction

When I was thirteen, I visited a dairy farm as part of a summer science program. The air was heavy with the sour-sweet smell of hay and manure. I remember being told how important dairy farms were for providing fresh products and supporting the local economy, but I was not listening to the guide.

Across a low fence, a cow stood still, bellowing. The sound was not sharp but drawn out and almost hoarse, a long, guttural cry that echoed through the metal shed. She did not appear to be physically injured, but when I looked over the gate, I noticed that the straw bedding beside her was empty.

"That's just how it works," the guide said.

"Each calf is separated from the mother early for its own good."

He tried to reassure me with a smile, but I stood there in borrowed boots, feeling disturbed and wondering when she would stop calling.

In the years since, I have often returned to that moment. I know that my reaction was emotional, but I do not think that makes it philosophically unimportant. Many people experience a similar discomfort when they see an animal suffering. Something about the suffering appears to demand a response before we have decided which ethical theory explains it. Yet the existence of that response does not tell us exactly why the suffering matters. Is the cow wronged because she has interests that deserve consideration? Does she possess rights that cannot be sacrificed for human convenience? Or does our discomfort reveal something deeper about the kind of moral agents we become when we learn to ignore suffering that benefits us?

These questions matter beyond individual acts of kindness. Contemporary societies depend heavily on institutions that use nonhuman animals: industrial agriculture, biomedical research, wildlife management, entertainment, and consumer production. At the same time, research in animal behavior and neuroscience has made it increasingly difficult to portray many animals as simple biological machines. A wide range of nonhuman animals exhibit pain behavior, social attachment, fear, memory, problem solving, and forms of emotional response. The ethical problem is therefore not merely that humans sometimes behave cruelly toward creatures we know nothing about. It is that our institutions increasingly recognize animal sentience while continuing to organize animal life mainly around human purposes.

Two influential approaches dominate modern philosophical debates about this

problem. The first is *interest-based ethics*, especially utilitarian theories associated with Jeremy Bentham and Peter Singer. These theories argue that an animal matters morally because it can suffer and therefore has interests that must be included in moral calculation. The second is *rights-based ethics*, most prominently defended by Tom Regan, which argues that at least some animals possess inherent value and cannot legitimately be treated merely as resources for others.

Both approaches identify something morally important, but neither is sufficient on its own. Interest-based ethics explains why animal suffering must count, yet its willingness to aggregate benefits and harms may permit serious violations of individuals when sufficiently large benefits are produced elsewhere. Rights-based ethics gives individuals stronger protection, yet it faces difficult questions about the basis and boundaries of rights and about conflicts between competing rights. I therefore argue for a qualified answer to the question posed in this essay: *we should treat nonhuman animals well because their interests are morally relevant and because rights can be necessary to protect those interests from sacrifice, but neither the language of interests nor the language of rights fully captures the moral problem.* Our treatment of animals also reflects how we draw the boundaries of moral visibility—whose suffering we notice, whose life we regard as expendable, and what kinds of exclusion we allow ourselves to normalize.

The final part of this essay develops this broader point through Foucault and Agamben. I do not claim that their political concepts can simply be transferred from human political history to animal ethics. Rather, their work helps illuminate a structural question already present in the debate: how institutions classify some forms of life as worthy of protection while making others available for use. This perspective does not replace interests or rights. It explains why both are necessary but incomplete.

2. Interest-Based Ethics: Why Suffering Matters

For much of Western intellectual history, animals were commonly placed below human beings within a moral and theological hierarchy. Augustine, for example, understood the created order in ways that gave human beings a distinctive moral status and did not treat animals as members of the same moral community.[1] Modern animal ethics represents a major challenge to this hierarchy because it asks whether species membership alone can justify giving human interests automatic priority.

[1] Augustine, *The City of God*, trans. Henry Bettenson (Penguin Classics, 2003).

The interest-based approach begins from a simple question: what feature of a being is morally relevant? Jeremy Bentham famously argued that the important question is not whether animals can reason or speak, but whether they can suffer.[2] Peter Singer develops this insight through the principle of equal consideration of interests.[3] If a being can experience pain, frustration, pleasure, or satisfaction, then what happens to that being matters from its own point of view. Its suffering must therefore enter moral consideration.

This does not mean that Singer believes animals and humans must always receive identical treatment. Equal consideration is not identical treatment. Different beings have different interests. A human being may have interests connected to long-term plans, political participation, education, or complex relationships that another animal does not possess in the same form. But where an interest is genuinely shared—most obviously the interest in avoiding severe pain—it cannot simply be discounted because the sufferer belongs to another species.

This is the strongest part of the interest-based position. It gives us a clear reason why animal suffering matters without requiring animals to possess human intelligence, language, or rationality. If pain is bad when experienced by a human, species membership alone does not explain why comparable pain should become morally irrelevant when experienced by another sentient creature. Singer calls the unjustified privileging of human interest *speciesism*, comparing its basic logical structure with other cases in which morally irrelevant group membership is used to discount another being's interests.[4]

The interest-based approach is also attractive because it is sensitive to consequences. Animal ethics often involves difficult questions in which every available option produces some harm. Agriculture, medical research, wildlife management, and environmental policy cannot always be addressed by a simple prohibition. A theory that compares suffering and benefit can therefore respond to moral complexity rather than pretending that ethical life contains no trade-offs.

Yet this strength also creates the theory's central weakness. If moral value is calculated through the aggregation of interests, then severe harm to one individual can in principle be justified by sufficiently large benefits to others. Utilitarianism does not regard

---

[2] Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart (Clarendon Press, 1996), chap. 17, sec. 1, para. 4n.

[3] Peter Singer, *Animal Liberation*, updated ed. (Harper Perennial, 2009).

[4] Singer, *Animal Liberation*.

suffering as morally irrelevant, but neither does it ordinarily treat the individual who suffers as inviolable.

Imagine that scientists discover a treatment that would permanently eliminate seasonal allergies for every human being. Billions of people would experience a real, though relatively minor, improvement in their lives. Suppose, however, that producing the treatment required the deliberate and painful killing of one kitten. If every benefit and harm were simply added together, the enormous number of small human benefits might outweigh the intense suffering of the single animal.

The example is intentionally extreme because it isolates the problem. Many people would still regard the act as wrong even if the total quantity of discomfort prevented was greater than the suffering caused. Their objection would not necessarily be that consequences never matter. Instead, they might think that there is something morally objectionable about treating one sentient being entirely as a container of costs that can be cancelled by benefits experienced by others.

This reveals a distinction between *having an interest* and *having a claim against others*. Interest-based ethics is very strong at explaining why the kitten's pain counts. It is less successful at explaining why there may be limits on what others are permitted to do to the kitten even when violating those limits produces a larger total benefit. The problem is therefore not that utilitarianism ignores animals. In fact, utilitarian philosophers have been among the most influential defenders of animal welfare. The difficulty is that a theory based on aggregate welfare may not provide sufficiently strong protection for individuals when their interests conflict with the interests of many others.

This is where rights-based theories become attractive.

3. Rights-Based Ethics: Why Individuals May Not Be Sacrificed

Rights-based ethics begins from a different moral intuition. Some things should not be done to an individual merely because doing them would benefit others. Rights place limits on the kinds of trade-offs that moral calculation may permit.

Tom Regan develops one of the most influential rights-based theories of animal ethics in *The Case for Animal Rights*. Regan argues that beings who are *subjects-of-a-life* possess inherent value. A subject-of-a-life does not merely exist biologically but has a life that can go better or worse for that individual. Such beings possess some combination of perception,

memory, desire, emotional life, preferences, and a sense of their own experiences over time.[5] Because their lives matter to them, they should not be treated merely as resources whose value depends entirely on their usefulness to others.

The rights-based approach directly addresses the problem raised by utilitarian aggregation. If an animal possesses a right not to be subjected to severe unnecessary harm, then the harm cannot simply disappear morally because enough other individuals benefit from it. Rights establish a boundary around the individual. They express the idea that moral subjects are not interchangeable units in a calculation of total welfare.

This protection is especially important in animal ethics because animals are usually much weaker than the humans and institutions that control them. A purely aggregate system can repeatedly place the burden of human benefits on creatures who have no political power to resist. Rights provide a language for explaining why some uses of animals may be impermissible even when they are profitable, convenient, or socially accepted.

However, rights-based theories generate problems of their own. The first concerns the basis of rights. If rights depend on being a "subject-of-a-life," philosophers must explain which capacities are sufficient and why. The problem is not simply that some animals obviously fall outside Regan's category. The deeper difficulty is that cognitive and emotional capacities exist on continua. Different species possess different forms of memory, social awareness, perception, planning, and sentience. Any threshold risks producing difficult borderline cases.

This does not refute rights theory, but it shows that rights cannot be assigned merely by announcing a category. If the threshold is set too high, some beings capable of substantial suffering may receive inadequate protection. If it is set very low, the theory must explain how rights operate across extremely different forms of life. Interest-based ethics handles this problem more easily because sentience can generate moral consideration even where the richer capacities associated with strong individual rights remain uncertain.

The second difficulty concerns conflicts among rights. Consider an ecosystem in which an invasive mammal species introduced by humans is rapidly destroying the population of an endangered native bird. If no intervention occurs, the native species may disappear. If the invasive animals possess strong rights against being killed, however, culling them may appear impermissible.

The problem is not that rights theories are literally incapable of discussing conflicts.

---

[5] Tom Regan, *The Case for Animal Rights* (University of California Press, 1983).

Rights can be ranked, limited, or interpreted in relation to other rights. But once these qualifications are introduced, the theory must explain how competing claims should be balanced. The more flexible the balancing becomes, the more the theory begins to reintroduce considerations of interests and consequences.

The two approaches therefore correct different moral dangers. Interest-based ethics prevents us from excluding animal suffering merely because the sufferer is not human. Rights-based ethics prevents us from treating individuals as endlessly exchangeable whenever aggregate welfare favors sacrifice. Yet each approach needs something supplied by the other. Interests explain why a being can be harmed. Rights explain why some harms should not simply be traded away.

For this reason, the most plausible answer is not that animals matter because they have interests *or* because they have rights. The two ideas address different parts of the moral problem. Sentient animals have interests because events can go better or worse for them. Where those interests are sufficiently important, rights can function as protections against their systematic sacrifice.

Still, even this combined position does not completely answer the question that disturbed me at the dairy farm. It can explain that the cow has an interest in maintaining social bonds and avoiding distress. It may also support rights limiting the ways she can be used. But a further question remains: *why are institutions able to make such interests and claims so easy for ordinary people not to see?*

That question moves beyond the distribution of moral rights and benefits toward the politics of moral attention.

4. Beyond Rights and Interests: Ethical Reflection and Moral Visibility

Neither interest-based nor rights-based ethics should therefore be discarded. The problem is not that the theories fail completely. It is that they begin after an important moral decision has already been made: the decision to recognize another being as someone whose interests or rights are worth considering.

Animal suffering is not hidden because humans have never developed moral principles capable of recognizing it. In many cases, the suffering is visible. What is striking is our ability to organize institutions in ways that make the suffering morally ordinary. Modern agriculture, research, consumption, and environmental management often require human beings to classify animal lives according to their usefulness. The ethical question is therefore

not only what animals possess but also how human institutions make some lives visible as moral subjects and others visible mainly as resources.

Michel Foucault's concept of *biopolitics* is helpful here. In *The History of Sexuality*, Foucault describes a form of modern power concerned not simply with prohibiting actions but with managing life: populations are measured, classified, regulated, and organized through institutions.[6] His primary subject is human political life, and the concept should not be casually transferred to animals as though the political histories were identical. Nevertheless, the framework raises a useful question for animal ethics. Modern institutions also manage animal populations on an enormous scale. Animals are bred, categorized, transported, confined, medically altered, reproduced, and killed according to systems designed around human objectives.

The relevance of Foucault is therefore structural rather than analogical. His work draws attention to the fact that power does not operate only through isolated acts of cruelty. It also works through systems that define what kinds of life can be managed and for what purposes. The moral difficulty of industrial animal use may therefore lie not only in individual harmful actions but in the normalization of a relationship in which an entire category of sentient life is administered primarily as a resource.

Giorgio Agamben develops a more radical vocabulary through the concept of *bare life*. In *Homo Sacer*, Agamben examines the political possibility of reducing a human being to biological existence while removing ordinary political protection. His central example, *homo sacer*, is a figure in Roman law who could be killed without the act counting as ordinary homicide.[7] Again, this concept comes from a specific analysis of human sovereignty and should not simply be equated with the condition of farmed animals.

Yet Agamben helps clarify a moral structure that matters here: the distinction between life that receives institutional protection and life that can be exposed to harm without being recognized in the same moral or political terms. Many animals in modern systems are intensely regulated precisely because they are valuable as biological organisms. Their reproduction, nutrition, health, movement, and death may all be carefully managed. But this management does not necessarily imply recognition of them as subjects whose own interests

---

[6] Michel Foucault, *The History of Sexuality*, vol. 1, *An Introduction*, trans. Robert Hurley (Vintage Books, 1990).

[7] Giorgio Agamben, *Homo Sacer: Sovereign Power and Bare Life*, trans. Daniel Heller-Roazen (Stanford University Press, 1998).

constrain the purposes of the institution.

The paradox is important. An animal can be highly visible economically and biologically while remaining morally marginal. The dairy cow is valuable enough for every aspect of her reproduction to be controlled, yet her distress at separation can be treated as an ordinary cost of production. Her life is not ignored in a literal sense; it is monitored carefully. What is limited is the moral significance assigned to her experience.

This reframes the question of animal ethics. Instead of asking only, "What rights does this animal possess?" or "How much suffering does this practice produce?" we can also ask:

What kind of moral community allows some forms of suffering to become structurally easy to disregard?

The question does not replace rights or interests. It explains why they may fail to influence practice even when we intellectually acknowledge them. A society can accept that animals are sentient and still design institutions in which their interests routinely lose against minor human benefits. It can create welfare rules while continuing to treat the animal principally as an economic unit. Moral theory therefore needs not only principles for resolving conflicts but also reflection on how the conflicts themselves are framed.

Agamben's work also suggests caution about the habit of exclusion. Here it is essential not to claim that animal exploitation is equivalent to the political persecution of refugees, colonized peoples, disabled persons, or other marginalized human groups. Such comparisons can erase important historical differences and can themselves become ethically careless. The more defensible point is structural: moral communities repeatedly rely on boundaries that determine whose experiences are taken seriously. Learning to regard vulnerability outside a favored category as morally irrelevant can become a dangerous habit of perception.

The lesson is therefore not that the same history repeats identically across species. It is that moral exclusion requires justification. If we become accustomed to treating membership in a category as sufficient reason for ignoring suffering, we weaken the demand that power explain why some lives receive less protection than others.

This insight returns us to interests and rights. Interests identify what is at stake for the animal. Rights place limits on what powerful actors may do with those interests. Ethical reflection on exclusion asks why those interests and rights become visible or invisible in the first place.

The three perspectives are therefore complementary rather than mutually exclusive.

5. Ethical Attention as Self-Definition

If this argument is correct, treating animals well is not only a response to facts about animals. It is also an expression of the kind of moral agents we choose to become.

This does not mean that morality should be reduced to human character. The cow's suffering matters because it is her suffering, not merely because cruelty makes humans unattractive. An ethical theory focused only on what animal treatment reveals about us would risk making animals instruments of human moral development once again. The point must therefore be stated more carefully.

Animal interests provide an independent reason for concern. Rights may provide strong protections for those interests. But *attention* determines whether we are willing to recognize those reasons when doing so is inconvenient.

The experience at the dairy farm mattered to me because the cow interrupted the ordinary description of what was happening. The guide presented separation as part of an efficient agricultural process. The cow's response made another description possible: a social animal had lost contact with her offspring and was distressed by that loss. Neither description was linguistically impossible. The moral question concerned which description I was willing to see as relevant.

Ethical attention therefore precedes some forms of moral calculation. Before we can weigh an interest, we must recognize that there is an interest to weigh. Before we can defend a right, we must regard the being who possesses it as the kind of individual whose claim deserves to be heard.

This does not produce a complete moral theory, and it should not pretend to. Attention alone cannot tell us whether invasive animals should be culled, which forms of biomedical research are justified, or exactly which rights different species should possess. These difficult questions still require principles, evidence, and judgment.

What attention does is prevent those principles from beginning with an invisible subject.

It reminds us that the beings affected by our institutions are not merely variables in economic, scientific, or ecological systems. They are living creatures for whom at least some things can go better or worse.

The ethical task is therefore both theoretical and reflective. We need interests because suffering matters. We need rights because suffering should not always be sacrificed to aggregate benefit. And we need attention because neither interests nor rights can protect

beings whose claims we train ourselves not to notice.

6. Conclusion

Should we treat nonhuman animals well because they have rights, interests, neither, or both? The most defensible answer is *both, but neither is sufficient by itself*.

Interest-based ethics provides the most basic reason why animals enter the moral community at all. Sentient animals can experience pain, fear, frustration, comfort, and satisfaction. What happens to them therefore matters to them. Their suffering cannot be dismissed simply because they belong to another species.

Yet interests alone are not enough. If every interest can always be aggregated with every other interest, then serious harm to vulnerable individuals may be justified whenever a sufficiently large number of others receive smaller benefits. Rights respond to this problem by expressing the idea that individuals possess claims that cannot always be exchanged for aggregate welfare.

Rights, however, also depend on judgments about which beings possess them, how strong different rights are, and how conflicts between rights should be resolved. The theory therefore needs continuing attention to interests, consequences, and differences among forms of life. Rights and interests should not be understood as competing explanations from which we must choose one. They protect different dimensions of moral standing.

But the debate remains incomplete if it stops there. Foucault and Agamben help reveal another dimension of the problem: institutions do not merely respond to already recognized moral subjects. They also shape which forms of life become visible as subjects at all. Modern societies can acknowledge animal sentience while continuing to organize billions of animal lives principally according to human purposes. An animal may be carefully managed as a biological and economic organism while its own experience remains morally secondary.

This is why the problem of animal ethics is also a problem of moral attention. Treating animals well requires more than finding the correct formula for balancing harms or assigning rights. It requires a willingness to notice the beings whose interests our ordinary practices make easy to ignore.

I therefore return to the cow I saw when I was thirteen. My sadness at the time did not amount to a philosophical argument. It did not tell me whether the cow possessed rights in Regan's sense or how her interests should be weighed against those of farmers and consumers. But the reaction contained something morally prior to those arguments: a refusal,

however incomplete, to regard her distress as merely part of "how it works."

Philosophy should do more than preserve that intuition, because intuitions can be inconsistent and mistaken. Interest-based ethics explains why the suffering I witnessed mattered. Rights-based ethics explains why the animal should not automatically be sacrificed whenever human convenience points in the opposite direction. Reflection on moral exclusion explains why an institution can make both facts difficult to see.

We therefore treat animals well not only because they have interests or because some possess rights, but because taking those interests and rights seriously is part of refusing an unjustified boundary between lives that count and lives that can simply be used.

The question of animal ethics ultimately returns to a question about moral responsibility. Power over another form of life creates the possibility of using it, but it also creates an obligation to justify that use. The greater our ability to control animal life becomes, the less convincing ignorance becomes as an excuse.

Perhaps this is what stayed with me from the dairy farm. The cow could not explain her claim in the language of rights or utility. She could only continue calling. The ethical work began when that sound became something I could no longer treat as morally irrelevant.

Bibliography

Agamben, Giorgio. *Homo Sacer: Sovereign Power and Bare Life*. Translated by Daniel Heller-Roazen. Stanford University Press, 1998.

Augustine. *The City of God*. Translated by Henry Bettenson. Penguin Classics, 2003.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Edited by J. H. Burns and H. L. A. Hart. Clarendon Press, 1996.

Foucault, Michel. *The History of Sexuality*. Vol. 1, *An Introduction*. Translated by Robert Hurley. Vintage Books, 1990.

Regan, Tom. *The Case for Animal Rights*. University of California Press, 1983.

Singer, Peter. *Animal Liberation*. Updated ed. Harper Perennial, 2009.

Review & Critique

Two Junior Reviews on Beauvoir and Woolf

Xichen Liu
Grade 10

Wency Xu
Grade 9

Editors' Introduction

This essay asks how different writers present gender identity and inequality to their readers. Its first part sets out Beauvoir's and Butler's accounts of female gender identity with clarity and identifies precisely where the two disagree. Its second part turns to Woolf, treating fiction, myth and imagined scenarios as forms of thought experiment. Setting the abstract genealogy of the first part against the imagined narrative of the second is a creative choice, and the essay offers well-articulated interpretations of each author's aims.

The essay supports its interpretations mainly with brief quotations rather than sustained close reading of the primary texts. As a result, its readings are often stated more fully than they are demonstrated from the texts. This is most evident in two places where interpretation is especially contested: its assessment of Butler's reading of Beauvoir, which is itself an interpretation of Beauvoir's text, and its account of *A Room of One's Own.*

Editors' note: This piece brings together two reading reviews originally written by students in the Junior group. They have been lightly edited and paired for publication, with the authors' original arguments and voices preserved.

Simone de Beauvoir and Virginia Woolf approach the condition of women through very different forms of writing. Beauvoir turns to history and genealogy, asking how "woman" came to occupy the position of the Other and why women have struggled to constitute themselves as a collective subject. Woolf often proceeds through another route: fictionalized narrators, imagined institutions, symbolic scenes, and counterfactual lives.

The following two junior reviews approach these writers separately. The first reads Beauvoir alongside Judith Butler and asks how history produces gendered subjecthood and political agency. The second examines Woolf's use of imaginative narrative, with Plato's *Symposium* as a point of comparison, and asks what fiction can make visible that abstract argument alone cannot.

Read together, the two pieces are concerned with a related question: how can writing make structures that appear natural, inevitable, or invisible available for critical thought?

1. Becoming "Woman": Beauvoir, Butler, and the Possibility of a Collective "We"

A genealogical analysis explores how our ideas about identity develop within particular historical contexts instead of treating these categories as permanent, unchanging truths. In *The Second Sex*, Simone de Beauvoir uses this historical approach to explain women's long-established position as the "Other." Under patriarchal systems, women's identities are largely defined relative to men, who are treated as the default independent subjects.

Without a collective history written from their own perspective, women's historical agency has often been obscured or insufficiently recognized. Beauvoir traces this unequal arrangement through social, economic, and cultural institutions that have developed over generations.

Judith Butler's 1986 essay "Sex and Gender in Simone de Beauvoir's *Second Sex*" complicates this framework.[1] Beauvoir treats biological sex as a material condition upon which gendered social existence is constructed. Butler questions whether "sex" itself can ever be treated as a neutral, pre-social category, since the ways bodies are classified and understood are also shaped by historical language and social norms.

Beauvoir's historical analysis therefore identifies the structural obstacles that keep women from forming a united, self-determined "We," while Butler raises a further question:

[1] Judith Butler, "Sex and Gender in Simone de Beauvoir's *Second Sex*," *Yale French Studies*, no. 72 (1986): 35–49.

if the category "woman" is itself historically produced, how stable can the basis of such a collective identity really be?

1.1 A History Written from Elsewhere

Beauvoir does not try to identify one single moment when women's oppression began. Instead, she traces how social expectations, legal rules, economic arrangements, and cultural narratives gradually combined over centuries to situate women as the Other.

Her statement that women "have no past, no history, no religion of their own"[2] should not be understood to mean that women were literally absent from history. Rather, recorded history has overwhelmingly been produced from male perspectives. Men largely determined which events counted as historically significant, while women were frequently represented through their relationships to men—as wives, daughters, mothers, or partners—rather than as independent historical subjects.

One major strength of Beauvoir's genealogical method is that it challenges the assumption that women's subordinate social status is natural or inevitable. If the position of women changes across historical periods and is repeatedly reinforced through institutions and conventions, then that hierarchy is historically constituted rather than simply given by nature.

This leads to an important political problem. If women have shared a marginalized social position for so long, why have they found it so difficult to unite against it?

1.2 Why Is There No Unified "We"?

Beauvoir compares women with other oppressed groups, especially the proletariat, whose members may develop a shared class consciousness and organize collective resistance.

Women, however, occupy a different structural position.

First, women are dispersed throughout society. They have historically lived within individual households and family structures rather than in separate communities organized around a common female identity. Without sustained common institutions or spaces, it becomes more difficult to develop lasting solidarity.

Second, women are materially and socially situated within relations of dependence that can make identification with men advantageous. Security, economic status, family position, and social recognition may all be tied to male family members. Under these conditions, attachment to the interests of men is not simply a matter of women consciously

[2] Simone de Beauvoir, *The Second Sex*, trans. and ed. H. M. Parshley (Vintage Books, 1989), xxv.

"accepting" oppression. It is produced by the structure of their lives.

Third, reproduction does not itself create oppression, but patriarchal societies have repeatedly interpreted women's reproductive capacities as justification for restricting them to domestic roles. Biological difference becomes translated into social hierarchy.

Together, these conditions help explain why women do not automatically constitute themselves as a unified political collective simply because they share forms of subordination. A common position does not necessarily produce common consciousness.

For Beauvoir, then, the difficulty of forming a "We" is historical and structural. Women are not merely individuals who have failed to recognize an obvious collective interest. The social organization of their lives has repeatedly made such collective identification difficult.

1.3 Butler's Challenge

Butler's critique begins with one of Beauvoir's most famous formulations: "One is not born, but rather becomes, a woman."[3]

Beauvoir's formulation distinguishes the biological conditions associated with sex from the social process through which someone comes to occupy the position of "woman." Gender is not simply an innate destiny contained in the body.

But Butler pushes this argument further.

For Butler, it is not enough to say that a biologically given sex later receives a socially constructed gender. The category of sex is also interpreted through historically produced norms. What counts as a stable, meaningful sexed identity is not independent of culture and language.

This is where performativity becomes important. Gender does not result from individuals consciously deciding to "perform" femininity or masculinity. Rather, repeated social acts and norms produce the appearance that a stable gender identity existed prior to those acts in the first place.

This complicates Beauvoir's account of collective political consciousness.

Beauvoir's analysis presupposes a materially situated category of women whose shared social conditions can provide a basis for collective political consciousness. Butler questions whether such a category can be established independently of the norms that produce gendered subjects themselves.

---

[3] Beauvoir, *Second Sex*, 267.

Experiences of gender are also shaped by class, race, culture, sexuality, and other social conditions. There may therefore be no single universal experience of womanhood capable of representing everyone included within the category "women."

But Butler's critique does not simply invalidate Beauvoir's historical account of women as the Other.

Instead, it changes how we might understand that process.

The position of the Other is not merely imposed upon an already complete and self-evident category of female bodies. Gendered identities and their subordinate social meanings are produced together across history.

1.4 Two Accounts of Agency

The difference between Beauvoir and Butler also changes how political agency can be imagined.

Beauvoir draws attention to the historical and material conditions that make collective consciousness difficult and therefore treats collective awakening as politically significant. Liberation requires women to recognize and transform the structures that have positioned them as dependent Others.

Butler shifts attention toward the instability of those structures themselves. If gender identity is continually reproduced through norms and repeated acts, then agency may emerge when those norms are disrupted, reworked, or exposed as contingent.

These perspectives should not be reduced to the claim that Beauvoir believes women simply need to "become aware," while Butler believes individuals simply need to behave differently. Both accounts are more structural than that.

Their disagreement concerns where the stability of gendered identity comes from and, therefore, how that stability might be challenged.

Beauvoir's genealogy exposes the historical and material arrangements that have kept women from forming an autonomous "We." Butler complicates the possibility of such a unified subject by questioning whether "women" can ever be treated as a category existing prior to the norms that constitute gender.

Read together, they reveal two dimensions of gendered subjecthood: the historical structures that position women socially, and the processes through which the identity "woman" itself becomes intelligible.

2. Imagining Inequality: Woolf and the Philosophical Power of Fiction

Philosophical truths are often associated with explicit arguments and abstract reasoning. Fiction, myth, and imagined situations seem to work differently. Their events may never have happened, yet they can make an idea easier to understand, experience, and remember.

Virginia Woolf's *A Room of One's Own* repeatedly uses this imaginative method. Her fictionalized narrator, the invented space of Oxbridge, the contrast between meals at men's and women's colleges, and the counterfactual life of Judith Shakespeare allow her to present gender inequality not simply as a proposition but as something a reader can picture.

A useful comparison can be found in Plato's *Symposium*.[4] Aristophanes does not define love through an abstract proposition. He tells a myth in which human beings were once whole and were later divided, leaving each half searching for the other.

The devices are not identical. Aristophanes employs myth; Woolf uses fictionalized narration, symbolic contrasts, and counterfactual invention. Yet both show why imaginative writing can have philosophical force: it turns an abstract idea into an experience the reader can enter.

2.1 Making an Abstract Idea Visible

In Aristophanes' speech, the idea of incompleteness becomes bodily.

Instead of simply saying that human beings desire companionship because they feel incomplete, the myth imagines bodies literally divided into halves. The image makes incompleteness visible and almost tactile. Readers can imagine what it means to have lost a part of oneself and to search for it again.

The myth therefore does not close the question of love with a definition. It offers an imaginative model that can be accepted, challenged, or compared with the other accounts of love in the *Symposium*. A philosophical question becomes something the reader can picture and emotionally inhabit rather than merely an abstract proposition.

This is one reason stories can be persuasive. They provide detail and structure. They give abstract concepts a shape.

But the same method is especially important in Woolf because her subject is not only an emotion such as love. It is institutional inequality.

---

[4] Plato, *Symposium*, 189c–193d.

2.2 Mary Beton and Oxbridge

A Room of One's Own employs a fictional narrative framework through its first-person speaker. Mary Beton is not simply identical with Virginia Woolf herself; she functions as a constructed perspective through which the experience of a female intellectual can be explored.

She moves through the fictional university town of Oxbridge, a composite that evokes Oxford and Cambridge. Through her movement, inequality is revealed in food, money, buildings, libraries, education, and physical space.

This matters because inequality can sound abstract when described only as a social structure.

Woolf instead makes it material.

The contrasting meals at men's and women's colleges turn differences in resources into something bodily and ordinary. Questions about institutional inequality become questions about what one eats, where one can enter, what books one can access, and whether one has enough money and privacy to think.

The famous requirement of "five hundred a year and a room with a lock on the door"[5] works in the same way. Money and private space are concrete, familiar things. Through them, Woolf shows that intellectual freedom is not produced by talent alone. It depends upon material conditions.

A personal anecdote might suggest that one individual woman happened to encounter disadvantages. Woolf's constructed setting allows her to present those disadvantages as part of an institutional structure.

The fictional framework therefore does not make the argument less serious. It gives the argument a world in which its consequences become visible.

2.3 Judith Shakespeare and What History Cannot Record

The example becomes even more striking in Woolf's invention of Judith Shakespeare.

Judith is imagined as Shakespeare's equally gifted sister. She possesses creative ability, but she does not receive the same education, freedom, mobility, or opportunity to write.

Her importance lies precisely in the fact that she is invented.

History can record the writers whose work survives. It can tell us whose plays were

[5] Virginia Woolf, *A Room of One's Own* (Harcourt Brace Jovanovich, 1989), chap. 6.

performed, whose books were published, and whose names entered literary history. But it has much more difficulty recording unrealized possibility.

It cannot easily show us the writers whose talent never had the conditions required to become visible.

Judith Shakespeare gives that absence a human form.

Woolf's point is not that this particular fictional woman literally existed and experienced every event described. The counterfactual situation allows the reader to ask what might happen when talent exists but the social conditions necessary for its development do not.

Historical absence can tell us that women were excluded from literary institutions. Fiction can make readers imagine the possibilities destroyed by that exclusion.

In this sense, imaginative writing can do something that a list of historical facts cannot.

It makes absence imaginable.

2.4 Why Not Simply State the Argument?

Woolf could simply have argued that female writers require economic independence, educational access, and private space.

But imaginative narrative changes how the reader encounters that claim.

When the argument appears through meals, libraries, rooms, and an invented sister of Shakespeare, the reader does not only receive a conclusion. The reader is asked to imagine the situations from which that conclusion emerges.

Narrative can therefore invite empathy and voluntary reflection. It can make philosophical and social questions feel less like statements imposed upon the reader and more like problems the reader has entered.

The imaginative example does not necessarily prove the argument, however.

That distinction matters.

A story may make a claim vivid without establishing that the claim is universally true. A fictional situation can reveal a structure without representing every person affected by that structure.

This becomes one of the limitations of Woolf's method.

2.5 The Limits of Fictional Representation

Fiction and allegory can be ambiguous because readers approach them with different experiences and historical backgrounds.

Even an image as famous as "five hundred a year and a room with a lock on the door"[6] may not be immediately understood by a reader living in another place or period. The problem Woolf describes may persist, but its form can change.

Fiction can also simplify complicated realities.

Woolf's imagined examples expose important structures of gender inequality, but no single narrative can account for every woman's experience. Historical realities differ by period, class, culture, and circumstance. Individual examples cannot become universal descriptions.

The same limitation appears in Aristophanes' myth.

The image of human beings searching for their lost halves gives love a memorable structure, but actual relationships are far more complicated. Love may involve desire, respect, independence, conflict, care, compromise, or forms of attachment that the myth cannot contain.

The problem, then, is not simply that fictional narratives are "false."

Their strength and their weakness come from the same source: they organize complexity into a particular image.

That image can illuminate one aspect of reality while leaving others outside the frame.

2.6 Fiction and Argument Together

Fictional narrative and philosophical argument are therefore not opposing modes of thought.

Narrative supplies images, experiences, and counterfactual situations through which abstract arguments become intelligible. Argument provides the conceptual structure through which those narratives can be interpreted and challenged.

In Plato, a myth turns incompleteness into a divided body.

In Woolf, a meal makes institutional inequality visible; a room turns intellectual independence into physical space; and Judith Shakespeare gives unrealized female possibility a name and a life.

These examples are persuasive not because imagination replaces evidence or

[6] Woolf, *Room of One's Own*, chap. 6.

reasoning. They are persuasive because they allow readers to understand what an abstract claim might mean within human experience.

The best use of narrative therefore does not ask the reader to choose emotion over reason.

It allows the two to work together.

3. Afterword: Two Ways of Making the Invisible Visible

Beauvoir and Woolf do not use the same method.

Beauvoir's genealogy asks how a social category becomes historically constituted. Her analysis makes visible the institutions, dependencies, and inherited narratives that have positioned women as the Other. Butler's critique then asks whether the category on which that history depends is itself more unstable than Beauvoir sometimes appears to assume.

Woolf works through a different form. Rather than tracing a genealogy, she repeatedly gives structures a scene: a meal, a library, a room, an imagined sister.

Yet both approaches resist treating the condition of women as self-evident or natural.

Beauvoir asks us to see a history where hierarchy once looked inevitable.

Woolf asks us to see a structure where inequality might otherwise remain abstract.

One makes the historical construction of gender available for analysis; the other makes its lived consequences available to imagination.

For two young readers approaching these texts today, perhaps this is part of what continues to make both writers compelling. They do not simply tell us what to think about women's lives. They give us different methods for asking how a social reality came to appear natural—and how writing might make it visible again.

Bibliography

Beauvoir, Simone de. *The Second Sex*. Translated and edited by H. M. Parshley. Vintage Books, 1989.

Butler, Judith. "Sex and Gender in Simone de Beauvoir's *Second Sex*." *Yale French Studies*, no. 72 (1986): 35–49.

Woolf, Virginia. *A Room of One's Own*. Harcourt Brace Jovanovich, 1989.

Review & Critique

A Home at the End of the World (2004): It's Just Love, Man. It's Nothing to Fear.

Olivia Liu
Grade 10

Editors' Introduction

This essay asks the interesting question of whether families founded on same-sex romantic love can be sustained. It answers through a close reading of the film *A Home at the End of the World*, set briefly against an episode of *Law & Order: SVU*. Its most substantial contribution is its analysis of "home" as a contested category. It innovatively reads Bobby as a figure onto whom the other characters project their desires, showing that the three-person household rests on asymmetrical forms of attachment. From this, the essay distinguishes romantic love from the broader affective bonds that can sustain a family.

The essay works primarily with narrative and dialogue rather than with cinematography, and it builds its argument through close reading rather than through queer theory or family studies. Its conclusions therefore speak most directly to the works it examines, while pointing toward the broader question of how families are constituted.

> Man, like, this whole big, beautiful, noisy world . . . . . and everything that can happen.
>
> —*A Home at the End of the World*[1]

A Home at the End of the World (2004), directed by Michael Mayer and based on the acclaimed 1990 novel by legendary author Michael Cunningham, with a screenplay he adapted himself, stars Colin Farrell as Bobby, Dallas Roberts as Jonathan, and Robin Wright as Clare. It remains a relatively obscure arthouse film that deserves a wider audience. Technically speaking, the film's execution and narrative pacing feel slightly unpolished, but the intrinsic narrative of the original book is rich enough that, despite its technical flaws, I still thoroughly appreciate it. While it is often tagged as a "queer film," I don't see it that way; queer desire is simply one facet of the vast, complex expressions of "love" this film explores.

The soundtrack captures the distinct musical essence of the late 1970s and 1980s, featuring folk rock, new wave, synth-pop, and blue-eyed soul. While it might sound a bit dated to modern ears, its strong period atmosphere aligns seamlessly with the film's mood.

True to its title, the concept of "home" serves as the central thread weaving the narrative together. As a young boy, Bobby loses his own family and moves in with Jonathan, becoming part of his household. Later in life, Bobby, Jonathan, and Clare form an unconventional family of their own and open a restaurant named "Home." By the end, this home narrows down to just two people, and perhaps soon, only one.

"It's just love, man. It's nothing to fear." These words are spoken at the beginning of the film by Bobby's older brother, Carlton, to the nine-year-old Bobby. High on hallucinogens or intoxicated, Carlton fails to notice the closed glass patio door in the dark, crashes through it, and suffers a fatal accident as glass shards pierce his carotid artery. It is the only absurd, wild, and youthful death in the entire story, while other characters merely die of illness or old age, marking the exact moment Bobby first experiences the profound loss of "home." Because Carlton's sudden death happens at such a formative age, Bobby's foundational sense of security is shattered before it can fully mature. He spends the rest of his life quietly seeking someone to anchor him in that lost safety, subtly laying the groundwork for how crucial the archetype of the "brother" will become to his emotional world.

Through Carlton, Bobby is exposed to marijuana at a young age, later encouraging both Jonathan and even Jonathan's mother, Alice, to try it. This suggests that under his

[1] *A Home at the End of the World*, directed by Michael Mayer, screenplay by Michael Cunningham (Warner Independent Pictures, 2004). All dialogue quoted in this review is from this film; later quotations from the film are not cited separately.

brother's influence, Bobby inherits a free-spirited, bohemian, almost hippie-like aura. With his long hair, he appears to be someone completely unbound by societal rules. Yet paradoxically, Bobby is far less proactive than he looks. Until he falls in love with Clare, he remains a virgin.

In his youth, Jonathan wears braces and glasses, coming across as somewhat awkward. He later attends NYU, growing into a young urbanite who better fits societal expectations. Yet within their relationship, Jonathan is consistently the more proactive one, whether initiating closeness with Bobby as a teenager or later openly embracing his sexuality and declaring he is the kind of person who needs a tattoo. Jonathan's active pursuit of passion stems from a deep need to escape the repressed quiet of his childhood home. Seeking self-definition in New York City, his willingness to embrace intimacy in a high-risk era shows that his search for connection was always urgent. While contracting HIV is by no means synonymous with promiscuity, in an era lacking awareness and protection, it subtly reflects Jonathan's more liberated attitude toward sex as he faces an HIV diagnosis.

In contrast, adult Bobby, who once seemed the more rebellious child, becomes the most innocent, least driven by physical desire among the three, always remaining optimistic. He seems capable of adapting to any environment, leading Clare to remark: "I bet you could live in the Sahara." Although a *Village Voice* review criticized that the movie shrank the character of Bobby to "a monosyllabic man-child with a puppy-dog stare,"[2] I still believe that in such a short movie telling a rather rich story, Colin Farrell's facial expressions don't make the characteristics of an innocent man feel weird.

Clare represents another approach to searching for a home. Having escaped a marriage with a violent ex-husband, she carries a distinct 80s punk ethos. In reality, most punks never lived as wildly as Sid Vicious (the Sex Pistols bassist who died of a heroin overdose after stabbing his girlfriend, Nancy). More often, they were simply rebelling against old lifestyles and crying out for love and peace, all while struggling to find a stable sense of belonging themselves. Clare seems like the most experienced person of the trio, yet she longs for a home more desperately than anyone. To maintain that sense of belonging, she helps create a family structure outside traditional marriage, only to realize that while the form of a family can be redefined, human emotional inclinations cannot be engineered. She believes that throwing away conventional rules will shield her from traditional heartbreak, yet

[2] Jorge Morales, "It's a Small World: Farrell's Man-Child Cut Down to Size," *Village Voice*, July 13, 2004, https://www.villagevoice.com/its-a-small-world-farrells-man-child-cut-down-to-size/.

emotional realities cannot simply be managed by design. The balance eventually broke down.

The scene where the three return to visit Alice after Jonathan's father passes away left a deep impression on me. In that moment, Jonathan says to Bobby in anger that if Bobby wants his mother, he can just have her. Throughout the story, everyone seems irresistibly drawn to Bobby: Alice, Clare, and even Jonathan himself.

It is here that I realized what makes Bobby so unique: he rarely demands or pursues anything actively, yet he has a quiet ability to let others project their desires onto him. Alice sees in him a lost youth and freedom; Jonathan sees a vitality he lacks; Clare sees the possibility of building a home. Having lost his original family so young, Bobby survived by becoming endlessly adaptable, offering unconditional warmth without demanding anything in return. Because he projects no rigid ego, Bobby functions as a vessel, holding everyone's ideal vision of life. And this passive charm is likely why, despite losing everything repeatedly, he manages to retain a quiet inner peace.

Jonathan, Bobby, and Clare all care, or at least they once did, for one another deeply on a spiritual level, but their love is neither identical in nature nor equal in intensity or purpose. That is why when Clare eventually asks Bobby, "Bobby. Do you wanna come with us?" and Bobby replies, "No," while looking at Jonathan, his decision carries immense weight.

Bobby tries to frame his past romantic connection with Jonathan as a piece of childhood confusion, yet he still says, "A little kiss between brothers." The line feels like both a denial and an admission, an attempt to categorize their past as simple adolescent intimacy, while being unable to deny Jonathan's singular place in his life.

This highlights the film's most painful yet tender reality: the three of them did create a real home together, but it was never built on equal footings of love. Someone loved more, someone needed more, and someone was simply searching for a place where they would never lose anything again.

What Bobby chooses to stay behind for may not be romantic love for another man, but rather the closest thing to a "home" he has ever known. As he tells Clare when she insists nothing could break him: "I couldn't be alone." Together, they scatter Jonathan's father's ashes on the hillside. By the end of the film, we return to the place where Bobby once took hallucinogens with his brother, only this time, it is young Jonathan sitting beside young Bobby. Throughout their complex emotional journey, Jonathan had stepped into the role of Bobby's older brother, the one who once told him, "It's just love . . . nothing to fear." In the end, they remain together in a place uniquely their own: *a home at the end of the world.*

What truly sustains a home? Is it romantic love?

The first time I encountered a different answer to this question was in an episode of *Law & Order: SVU*. It featured a lavender marriage where both partners were fully aware of each other's true sexual orientations, yet chose to stay together, not out of romantic passion, but out of deep familial affection, trust, and companionship.

Not everyone in this world experiences romantic love, nor does everyone seek it. The film portrays a triangular relationship that ultimately falls apart, but it is undeniable that this unconventional family *did* work for a time. In the real world, such arrangements can and do succeed. They gave each other what they desperately needed and, together, built something that was genuinely a home.

We, varied as we are, might chase traditional romance, or find ourselves clinging to unconventional forms of belonging; we might stumble through crises of trust, or be afraid to say love out loud; sometimes, we don't end up with romantic love at all, and that is fine. What truly matters is the willingness to open ourselves up to human connection, in whatever form it takes. After all, it's just love, man. It's nothing to fear.

Bibliography

Mayer, Michael, dir. *A Home at the End of the World*. Screenplay by Michael Cunningham. Warner Independent Pictures, 2004.

Morales, Jorge. "It's a Small World: Farrell's Man-Child Cut Down to Size." *Village Voice*, July 13, 2004. https://www.villagevoice.com/its-a-small-world-farrells-man-child-cut-down-to-size/.

Review & Critique

On the Ending Scene of *Pygmalion* and *My Fair Lady*

Fanxiao Sun
Grade 12

Editors' Introduction

This essay traces the Pygmalion narrative from Greek myth through Shaw's play to the 1964 film *My Fair Lady*. It incisively argues that the adaptation reverses the balance of power between Higgins and Eliza. The argument rests on close comparative reading, notably of Higgins's exposed vulnerability in "I've Grown Accustomed to Her Face" and of Eliza's return as an act of choice. On this basis, the essay delivers a novel reading of the film's ending as an ambiguous compromise, neither straightforwardly conservative nor conventionally romantic, and relates that ambiguity to the cultural tensions of the 1960s.

The essay's strength lies in its textual account of *how* the ending was transformed. Its account of *why* rests on broader claims about the studio system, Audrey Hepburn's star persona and postwar gender ideology, which it advances as interpretive context rather than establishing in detail. Other forces of the period, such as the emergence of Second Wave feminism, remain open as possible explanations.

In ancient Greek mythology, Pygmalion, king of Cyprus, was so disgusted by the depravity of women that he resolved to remain unmarried, yet he fell in love with an ivory statue of a woman that he had carved himself. Eventually, the goddess of love gave the statue life, and the two married. In 1912, the Irish playwright George Bernard Shaw used this myth as a blueprint to create the play *Pygmalion*,[1] transforming the artist's "creation" into a linguist professor's "remaking" of a lower-class flower girl. Half a century later, the play was first adapted into a Broadway musical, and then in 1964 into the classic musical film *My Fair Lady*.

However, the adaptation from play to film was by no means a simple change of medium. Shaw's original is a sharp social satire, whereas the film *My Fair Lady*, with its lavish songs and romantic ending, dilutes this social critique into an entertaining story of personal growth and love. In no part is this transformation more pronounced than in the ending, whereby the open ending where Eliza is suggested to have left Higgins for good and marry Freddy in the original play is rewritten to a romanticized version, with Eliza eventually returning after the quarrel.

This transformation is the product of two factors—one external, rooted in commercial imperatives and market expectations; and the other internal, operating within the narrative's own psychological and thematic framework. The former has been widely acknowledged in previous studies; the latter, however, reveals a more subtle and even subversive dimension to the film's apparently conventional closure.

My Fair Lady was produced at the height of the Hollywood studio system in the 1960s, and its creative decisions were first and foremost constrained by demands of commercial markets. As a musical, the film was bound by the conventions of the Hollywood musical style, a form that demanded emotional uplift and beautiful resolution. Twenty-one musical numbers, opulent costumes, and extravagant set designs collectively constructed an audience expectation of spectacle and romance. To conclude with Eliza's angry departure and Higgins's solitary dejection would have dissatisfied such expectations and jeopardized the film's commercial viability.

Equally significant was the logic of the star system. Audrey Hepburn, cast as Eliza, embodied the archetype of the "elegant commoner" that 1960s Hollywood adored. Under the regime of star-driven audience identification, viewers could not easily accept their beloved Hepburn ending alone and uncertain. The screen demanded a "homecoming" for its heroine,

[1] George Bernard Shaw, *Pygmalion*, ed. Dan H. Laurence (Penguin Books, 2003).

and the safest, most emotionally satisfying homecoming was always romantic closure—far more so than the pragmatic independence of marrying Freddy and opening a flower shop, as Shaw's epilogue had suggested.[2]

Beneath these commercial pressures lay a deeper ideological reason. In the postwar American context of economic prosperity, when women were increasingly tempted to return to household duties, the declaration of an independent female was a more provocative proposition than it had been in Shaw's 1910s Britain, where the "suffrage" movement was in its prime. The film's solution—allowing Eliza to storm out but ultimately return—was an act of ideological compromise, aligning with a long tradition of romantic comedy in which the heroine's anger is ultimately subsumed by love.

Yet to dismiss the film's ending as mere submission to commerce and conservatism would be to overlook what's embedded within the narrative itself. The adaptation, operating within its own internal logic, performed a remarkable act of transposition—shifting Shaw's social manifesto into a psychological drama about the inversion of power between creator and creation, god and mortal, Pygmalion and Galatea.

In the ancient tale, Pygmalion remains the unchanging creator, and Galatea's animation only confirms his artistic triumph and romantic desire; the union is completed without any transformation on his behalf. Shaw's play broke that by making Galatea refuse her creator, but the film takes a more circuitous route: it restores the physical reunion, yet undermines its mythological meaning by reversing the very power dynamics that the myth celebrated.

We witness the reverse transformation of Higgins from god to mortal. In Shaw's play, Higgins is the untouchable Olympian figure—arrogant, self-sufficient, and untouched by the experiment he conducts. The film, however, grants him a moment of vulnerability through the song "I've Grown Accustomed to Her Face." Alone in his study, he cannot work, cannot listen to the recording without anguish, and cannot escape the void left by Eliza's absence. This is a crucial innovation: the creator undergoes a metamorphosis as profound as the creation's. His "godlike" detachment crumbles, revealing a man undone by habit and dependency. Where the mythic Pygmalion only ever descends his throne to claim his beloved, the film's Higgins must first be dethroned—stripped of his intellectual sovereignty—before any reconciliation becomes possible. When Eliza finally returns, she confronts not the omniscient professor but an almost pitiable figure who has already been "conquered" by his

[2] Shaw, *Pygmalion*, "Sequel: What Happened Afterwards."

own need for her. Their reunion, therefore, is not a submission of the creation to the creator, but a meeting between two fundamentally altered beings—a dynamic that the original myth never entertained.

Then there is also Eliza's elevation from a creation to an autonomous human being. If Higgins is brought down from his divine pedestal, Eliza is simultaneously lifted from the status of a mere "artifact" to that of a full subject in her own right. In Shaw's play, this elevation is achieved through her final act of departure—she walks out to claim her independence. The film, however, accomplishes the same elevation through a more paradoxical gesture: her voluntary return. In the final scene, Higgins plays an old recording of Eliza's old pronunciation at their first meeting in the house. As the voice fills the room—the ghost of the "old Eliza"—the door opens, and the real Eliza stands before him, speaking after the unfinished dialogue, "I washed my hands and face before I come, I did."[3] In myth, Galatea's animation is a miracle granted by a goddess; she receives life and voice passively. Here, however, Eliza actively deploys what used to be regarded as her linguistic deficiency as a combating sword against her remaker. She is indeed returning to the same man and the same maltreatment, but she returns as one who has already proven she can leave—and who therefore stays by choice, not by necessity.

In this shift of focus from a resolute "goodbye" to a somewhat ambiguous "see you again," the film also signifies a more subtle proposition: an individual's independence is not necessarily emphasized through departure. When one has the option to leave yet chooses to return, and does so with full self-awareness, that very act constitutes a confirmation of one's own subjectivity. Shaw has Eliza declare her independence through walking out; the film, however, lets her declare it through coming back—"I am back, but I come back of my own accord."

Perhaps the most unexpected narrative choice is the setting of the finale: when Higgins asks out the famous line, "Eliza, where the devil are my slippers?" The film does not have Eliza erupt in outrage—as Shaw's play would do—but instead gives her a knowing, ambiguous smile. This gesture performs a quiet subversion: by having Eliza return, the film refuses the grandiosity of either tragic separation or triumphant independence. She returns not to a palace or a wedding chapel, but to the most mundane and imperfect of domestic arrangements—a man who will never stop being rude, a room that will never be tidy, a life

[3] *My Fair Lady*, directed by George Cukor (Warner Bros., 1964). All dialogue quoted in this essay is from this film; later quotations from the film are not cited separately.

that will never be mythic. In the original myth, Pygmalion and Galatea ascend to a divine marriage, a union blessed by Aphrodite herself; their story ends in legend. The film, by contrast, brings its protagonists down to earth, embedding them in the trivialities of shared existence—slippers, records, and unspoken resentments. Her return deflates the heroic postures of both the "gallant rebel" and the "redeemed woman." She comes back not because she has nowhere else to go, but because she chooses to—and that act of choosing, within the constraints of a messy everyday life, is its own quiet assertion of agency. The film, in effect, rewrites Shaw's manifesto into a situation comedy about two flawed individuals learning to coexist—a move that sacrifices the sharpness of social critique, but gains a more intimate, almost anti-mythical realism.

Taken together, it could be argued that the film's ending, far from a simple capitulation to the myth's romantic template, is a deliberate reworking of it. On the surface, Eliza returns, and Pygmalion seems to have won his Galatea—just as the ancient story promises. But the internal logic has hollowed out that victory: Higgins is no longer the confident sculptor, and Eliza is no longer the passive statue. The myth celebrates the creator's triumph; the film, read closely, celebrates the creation's ironic triumph over the creator. In this sense, the film does not abandon Shaw's subversive spirit; it translates it into a different form of expression.

Of course, this "creative paradox" comes at a cost. The ambiguity of the film's ending makes it easy for mainstream audiences to read it simply as "happy ever after," thereby dissolving whatever subversive potential it might contain. Yet it is precisely this ambiguity that renders the ending of *My Fair Lady* still worthy of sustained interpretation half a century later—for it satisfies the commercial demand for happy closure, while simultaneously planting beneath that surface a narrative seed of power reversal.

Thus, while commercial imperative demanded that Eliza come back, the narrative worked within that constraint to make her coming back mean something far more complex than commercial demand. The film gives us an ending that satisfies the crowd and haunts the attentive viewer in equal measure—a compromise, perhaps, but a remarkably fertile one, because it quietly transforms the ancient myth of divine creation into a modern fable about the mutual vulnerability that alone makes true equality possible.

Bibliography

Cukor, George, dir. *My Fair Lady*. Warner Bros., 1964.

Shaw, George Bernard. *Pygmalion*. Edited by Dan H. Laurence. Penguin Books, 2003.

Review & Critique

Lies Matter—The Punishment of Lying Politicians

Yiming Chen
Grade 12

Editors' Introduction

This essay takes up the longstanding problem of lying among public officials and argues that politicians should be held accountable for lies about public affairs. It moves across moral philosophy and political theory, setting out utilitarian and Kantian positions and the problem of "dirty hands," and it tests these frameworks against journalistic accounts and contemporary cases. It engages critically with established philosophers, at times arguing against them. Its most useful conceptual contribution is a clearly drawn distinction between public and private lying, which it applies effectively to the present-day political scene.

The essay's focus is normative: it argues that political lies should be sanctioned, while leaving the form of that sanction open. It also points toward, without pursuing at length, a further question: why lies about public affairs so often go unpunished when lies about private conduct do not.

1. Introduction: Deception, Democracy, Derogation

On January 28, 1998, in front of cameras of anxious journalists, the US President Bill Clinton made a statement that shocked the world. Commenting on the "rumor" of his sexual relationship with White House intern Monica Lewinsky, he asserted that he "did not have sexual relations with that woman, Miss Lewinsky."[1] However, after being hard-pressed by increasingly indisputable evidence, Clinton finally admitted his "'inappropriate' conduct with Lewinsky" on August 17, 1998[2] and confessed that he "lied under oath at a civil deposition."[3] The scandal led to an impeachment for perjury, a federal crime for lying under oath;[4] yet the impeachment failed, with Clinton promptly acquitted.[5] The failure of the impeachment rekindled disputes over a controversial political question: should politicians be punished for lying? If so, where should the limit of private and therefore unpunishable lies of statesmen, as differentiated from their punishable lies made in the public sphere, be rightfully drawn? This essay argues that within the contemporary legal framework, politicians lying on public affairs should be punished.

Certain conceptual definitions must be clarified before discussion. This essay adopts a narrow definition of "lying," distinguishing it from other camouflage techniques. Jurecic defines lying as "an act undertaken intentionally to obscure the truth and which therefore must be performed with a knowledge of the facts," thereby distinguishing it from "bullshit," an act "undertaken without any relationship to truth whatsoever."[6] We can achieve a more accurate definition by referencing Sandel's distinction between a lie and a misleading

---

[1] CBSN, "This Day in History: Bill Clinton Says 'I Did Not Have Sexual Relations with That Woman,'" YouTube video, January 28, 2016, https://www.youtube.com/watch?v=_aGbdni7QNs.

[2] Russell Riley, "The Clinton Impeachment and Its Fallout," Miller Center, 2025, https://millercenter.org/the-presidency/impeachment/clinton-impeachment-and-its-fallout.

[3] "Star-Bulletin Special: The Starr Report," Starbulletin.com, accessed July 5, 2025, https://archives.starbulletin.com/specials/starr/intro.

[4] Perjury Generally, 18 U.S.C. § 1621 (2024), https://uscode.house.gov/view.xhtml?req=%28title%3A18+section%3A1621+edition%3Aprelim%29.

[5] Riley, "Clinton Impeachment and Its Fallout."

[6] Quinta Jurecic, "On Bullshit and the Oath of Office: The 'LOL Nothing Matters' Presidency," *Lawfare*, November 23, 2016, https://www.lawfaremedia.org/article/bullshit-and-oath-office-lol-nothing-matters-presidency; Harry G. Frankfurt, *On Bullshit* (Princeton University Press, 2005), quoted in Jurecic, "On Bullshit and the Oath of Office."

statement, the latter being an "artful dodge" of partial truth instead of the former's "bald-faced lie."[7] A distinction between public and private lies must also be drawn. Within the liberal school's two prevalent methods of defining the private sphere—broadly defining it as issues "not governed by the public power of the state," or narrowly defining it as issues inside "a sphere of intimacy to which one might retreat in face of the pressures to conform within society"—this essay adopts the narrow definition to emphasize the difference between a politician's actions conducted as a private person and those as a public figure.[8]

Proper definitions of moralism and utilitarianism in the political perspective are also indispensable to our analysis. Here, this essay adopts Maynard's definition of moralism: "the view that proper practical normative judgements about politics . . . depend, in some sense, on moral judgements."[9] Meanwhile, Foot defines utilitarianism as the combination of the belief that "the rightness or goodness of actions . . . consists in their positive productive relationship to these states of affairs" and "the identification of the best state of affairs . . . in which there is most happiness, most pleasure, or the maximum satisfaction of desire."[10] Finally, discussions in this essay will be confined to democratic nations, as lying attacks "at the very heart of democracy and the whole project of enlightenment epistemology"[11] and makes "political discourse" impossible and democracy "in crisis."[12] Moreover, leaders of autocratic states with their de facto existence above the law cannot be spoken as legally punished or punishable: just as their multiple outrageous actions are unrestrained by laws, their lies are likewise hardly identified—at least in public—let alone penalized.

---

[7] Michael Sandel, "White Lies," *New Republic*, March 2, 1998, 10–11.

[8] Judith Squires, "Public and Private," in *Political Concepts*, ed. Richard Bellamy and Andrew Mason (Manchester University Press, 2018), 132, https://doi.org/10.7765/9781526137562.00015; Will Kymlicka, *Contemporary Political Philosophy: An Introduction*, 2nd ed. (Oxford University Press, 2002), 388–98.

[9] Jonathan Leader Maynard, "What Is Political Moralism?," *Topoi* 44, no. 1 (2025): 14, https://doi.org/10.1007/s11245-024-10060-9.

[10] Philippa Foot, "Utilitarianism and the Virtues," *Mind* 94, no. 374 (1985): 196, https://doi.org/10.1093/mind/xciv.374.196.

[11] James Pfiffner, "Trump's Lies Corrode Democracy," Brookings, April 13, 2018, https://www.brookings.edu/articles/trumps-lies-corrode-democracy/.

[12] David Johnson, "Why False News Endangers Democracy," *Psychology Today*, December 10, 2016, https://www.psychologytoday.com/us/blog/constructive-controversy/201612/why-false-news-endangers-democracy.

2. Justifications and Detriments of Political Liars

The discussion begins with the argument that lies are ethically detrimental in nature and therefore should receive moral punishments, whatever the circumstance might be. Sullivan, for instance, presses the necessity of punishing lying politicians by arguing that such persons as Clinton "demeans and threatens the entire system of government he is elected to defend."[13] A common utilitarian counterargument, of course, would emerge at this point with its justification of "white lies" that bring benefits. Sandel argues that Franklin Roosevelt's denial of his plans of entering the Second World War in 1940 was justified, since it was "for the sake of a just cause";[14] however, the statement that lying might generate positive utilitarian results cannot obscure the natural immorality of lies. Kant comments that the so-called white lies do not justify deceit by their self-proclaimed intentions but only "escape being punishable merely by accident," yet on the other hand "public justice can hold nothing against" telling the truth "whatever the unforeseen consequences might be."[15] In other words, despite that a lie and its consequence receive separate moral judgements and sometimes a lie can cause a positive result, the very fact of deceit should always be penalized. Adding to that, Sidgwick points out if "the lawfulness of benevolent deception in any case" was admitted, we could not "decide when and how far it is admissible by any clear method except that of utilitarianism," a criteria against "Common Sense" in this case.[16] Therefore, a "white lie" should not be exempted from punishment completely.

Some might also argue that lying as a practical tradition of politicians is within their administrative discretion despite its ethical-legal impropriety, and therefore should be exempted from punishment. For instance, Sartre emphasizes the "dirty hand" norm by using Communist leader Hoerderer's statement: "I have dirty hands right up to the elbows. I've plunged them in filth and blood. Do you think you can govern innocently?"[17] Walzer, however, suggests a moral rebuttal against the two counterarguments by stressing that the

---

[13] Andrew Sullivan, "Lies That Matter," *New Republic*, 1998, reproduced at UltraFaction, accessed July 5, 2025, https://www.ultrafaction.com/articles001/anti-clinton-sullivan-98.html.

[14] Sandel, "White Lies," 10–11.

[15] Immanuel Kant, "On a Supposed Right to Lie from Philanthropy," in *Practical Philosophy*, ed. Mary J. Gregor (Cambridge University Press, 1996), 605–16, https://doi.org/10.1017/CBO9780511813306.014.

[16] Henry Sidgwick, "The Classification of Duties—Veracity," in *The Methods of Ethics*, bk. 3, chap. 7 (Palgrave Macmillan, 1962), 312–19, https://doi.org/10.1007/978-1-349-81786-3_22.

[17] Jean-Paul Sartre, *Dirty Hands*, trans. Gerard Hopkins (Methuen Drama, 2001), 224.

status quo is maintained not because it is morally accepted, but for the fact that "there is rarely a Czarist executioner waiting in the wings for politicians with dirty hands," even if "most of us believe that those who govern us" are actually known by their "dirty hands."[18] This is because "we won't be able to do that [i.e. punish the lying politicians], however, without getting our own hands dirty."[19] With a firmly moral stance, he insists on punishing lying politicians and calls for public action to conduct the punishments. The Editorial Board of *New York Times* expressed a similar argument in 2016, stating that "in the absence of" political leaders who "care about the truth," "citizens who think for themselves will need to light the way."[20] Whether Walzer's argument is excessively idealistic or not, the current commitment of the civil society on punishment against lying politicians is indeed far from satisfactory.

As demonstrated by Sidgwick's "Common Sense" rebuttal above, the completely utilitarian perspective on this issue of lying politicians is unacceptable. Despite that the practicability of Walzer's staunchly moral attitude is still susceptible, three fundamental conclusions can at least be drawn from the rebuttals of Kant, Sidgwick, and Walzer: one, lying is by nature detrimental and thus should never be exempted from punishment, even for a politician with seemingly benevolent motivation; two, lying should always be punished from a theoretical and ethical perspective, whatever the reality of juridical-political practice is; three, the contemporary custom of not punishing deceitful statesmen in the political arena needs a change. Such conclusions, unfortunately, still fail to answer our question: should we adopt Walzer's perspective and determine that all lying politicians should be punished? Although the positive answer is indeed tempting, the societal effects of this idealistically moral stance must be considered: we must consider the aforementioned boundary between public and private lies. This essay argues that only the politicians who lie on affairs related to their public duty should be punished, while they should still enjoy the right to lie privately like any others.

Of course, the differentiation is unnecessary if the lying behavior is illegal in itself, such as perjury. Within legal practice, one consequence of encouraging punishment against politicians' private lies is the misdirection of public attention. Lewis, for instance, points out

---

[18] Michael Walzer, "Political Action: The Problem of Dirty Hands," *Philosophy & Public Affairs* 2, no. 2 (1973): 161, 168, https://doi.org/10.2307/2265064.

[19] Walzer, "Political Action," 180.

[20] "Truth and Lies in the Age of Trump," editorial, *New York Times*, December 10, 2016.

that prosecutors in the Clinton-Lewinsky Affair were actually investigating sex instead of lies: a dangerous trend of overemphasizing private scandals at the expense of investigative justice on public crimes should thus be noted.[21] The trend of prioritizing defaming scandals over real legal offences can also be extended to other private areas, and punishments against lies on private affairs would easily fuel such a trend if encouraged. This dangerous trend can lead to serious consequences: the zeal of gossiping individuals and media would end "the privacy of anyone who aspires to office,"[22] which entails an aggression against the politicians' human and civil rights. From a utilitarian perspective, the focus on private affairs would force the prosecutors to investigate private affairs more frequently and lead to impaired procedural justice, and encourage partisan witch-hunts based on whispering campaigns and vaguely accusations of immorality.

The common counterargument that deception on private affairs reflect the general moral fallacy of a politician would naturally emerge here. Sullivan, for instance, commented on the Clinton-Lewinsky Affair that "the lies about sex are not exceptions to the general rule; they are the rule" for Clinton.[23] However, despite that an exmarital sexual scandal such as Clinton's is indeed morally reprehensible, it is risky to punish such private outrages by circumventing the law. Legally, politicians should be punished only after they have expanded their lies to the public realm or extended their moral untrustworthiness into public commitments. Some would still argue that the politicians' lies almost always have impacts upon their capability of administrative duty: for instance, politicians covering up their diseases always suffer from a decline in competence of satisfying administrative responsibilities. Similarly, a politician may lie about an inappropriate personal handling with businesses, which later results in the corporation's successful extortion against his/her policymaking. However, in the first scenario, the decision to lie about a leader's health issues is usually a governmental consensus for security reasons that requires the agreement of representative legislators, rather than the personal will of the statesman alone. In the second scenario, the improper personal handling in itself entails illegality, the punishment of which needs not consult the various attitudes on private lies. Therefore, the potential public consequences of private lies still do not automatically justify punishments against the statesmen's private lies.

---

[21] Anthony Lewis, "Sex and Leadership," *New York Times*, February 23, 1998.

[22] Lewis, "Sex and Leadership."

[23] Sullivan, "Lies That Matter."

Unsurprisingly, Lewis isn't alone in defending politicians' lies in the private realm. The US First Amendment, in fact, also provides a legal framework for protecting lies in the private sector but not those related to public responsibilities. The First Amendment, stating that "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances,"[24] gives private lies a constitutional protection. However, regulations on speech can be justified if they meet a heavy burden, "which requires the government to prove the regulation furthers a compelling interest and is narrowly tailored to achieve that interest."[25] According to Steiner and Eduardo, interests are "compelling" when they are regarding "regulation vital to the protection of public health and safety, including the regulation of violent crime, the requirements of national security and military necessity, and respect for fundamental rights."[26] Interpreting the burden from the perspective of lying regulation, it authorizes the right to regulate and punish lies related to public affairs. Therefore, politicians' lies on private affairs should be protected, while those related to public duties should be punished.

3. Conclusion: Fine Balance Between Privacy and Punishment

The status quo of punishment against lying politicians is indeed far from satisfactory. Yet from a moral or utilitarian perspective, this essay does not encourage punishment against politicians' lies on the private sector. More than twenty years ago, Clinton lied under oath yet still avoided effective punishment, with the media paying "excessive focus on the president's sexual conduct" instead of the lies themselves.[27] To avoid drawing too much public attention to private affairs and wasting national resources, we must give a leeway for politicians to lie

---

[24] U.S. Const. amend. I, Constitution Annotated, Library of Congress, https://constitution.congress.gov/constitution/amendment-1/.

[25] *Reed v. Town of Gilbert*, 576 U.S. 155 (2015), First Amendment Library, Foundation for Individual Rights and Expression, accessed July 18, 2025, https://www.thefire.org/supreme-court/reed-v-town-gilbert; Angel Eduardo, "Why (Most) Lies Are Protected Speech, and Why They Should Stay That Way," Foundation for Individual Rights and Expression, June 29, 2023, https://www.thefire.org/news/why-most-lies-are-protected-speech-and-why-they-should-stay-way.

[26] Eduardo, "Why (Most) Lies Are Protected Speech."

[27] Leigh Husband Kimmel, "President Bill Clinton Denies Sexual Affair with a White House Intern," EBSCO Research Starters, EBSCO Information Services, 2022, https://www.ebsco.com/research-starters/political-science/president-bill-clinton-denies-sexual-affair-white-house-intern.

about private affairs. Yet to protect the democratic system and maintain trust, we must punish those lying about affairs relevant to his public responsibility to the people.

Bibliography

CBSN. "This Day in History: Bill Clinton Says 'I Did Not Have Sexual Relations with That Woman.'" YouTube video, January 28, 2016. https://www.youtube.com/watch?v=_aGbdni7QNs.

Eduardo, Angel. "Why (Most) Lies Are Protected Speech, and Why They Should Stay That Way." Foundation for Individual Rights and Expression, June 29, 2023. https://www.thefire.org/news/why-most-lies-are-protected-speech-and-why-they-should-stay-way.

Foot, Philippa. "Utilitarianism and the Virtues." *Mind* 94, no. 374 (1985): 196–209. https://doi.org/10.1093/mind/xciv.374.196.

Frankfurt, Harry G. *On Bullshit*. Princeton University Press, 2005.

Johnson, David. "Why False News Endangers Democracy." *Psychology Today*, December 10, 2016. https://www.psychologytoday.com/us/blog/constructive-controversy/201612/why-false-news-endangers-democracy.

Jurecic, Quinta. "On Bullshit and the Oath of Office: The 'LOL Nothing Matters' Presidency." *Lawfare*, November 23, 2016. https://www.lawfaremedia.org/article/bullshit-and-oath-office-lol-nothing-matters-presidency.

Kant, Immanuel. "On a Supposed Right to Lie from Philanthropy." In *Practical Philosophy*, edited by Mary J. Gregor. Cambridge University Press, 1996. https://doi.org/10.1017/CBO9780511813306.014.

Kimmel, Leigh Husband. "President Bill Clinton Denies Sexual Affair with a White House Intern." EBSCO Research Starters. EBSCO Information Services, 2022. https://www.ebsco.com/research-starters/political-science/president-bill-clinton-denies-sexual-affair-white-house-intern.

Kymlicka, Will. *Contemporary Political Philosophy: An Introduction*. 2nd ed. Oxford University Press, 2002.

Leader Maynard, Jonathan. "What Is Political Moralism?" *Topoi* 44, no. 1 (2025): 13–26. https://doi.org/10.1007/s11245-024-10060-9.

Lewis, Anthony. "Sex and Leadership." *New York Times*, February 23, 1998.

New York Times. "Truth and Lies in the Age of Trump." Editorial, December 10, 2016.

Pfiffner, James. "Trump's Lies Corrode Democracy." Brookings, April 13, 2018.

https://www.brookings.edu/articles/trumps-lies-corrode-democracy/.

Reed v. Town of Gilbert. 576 U.S. 155 (2015). First Amendment Library, Foundation for Individual Rights and Expression. Accessed July 18, 2025. https://www.thefire.org/supreme-court/reed-v-town-gilbert.

Riley, Russell. "The Clinton Impeachment and Its Fallout." Miller Center, 2025. https://millercenter.org/the-presidency/impeachment/clinton-impeachment-and-its-fallout.

Sandel, Michael. "White Lies." *New Republic*, March 2, 1998, 10–11.

Sartre, Jean-Paul. *Dirty Hands*. Translated by Gerard Hopkins. Methuen Drama, 2001.

Sidgwick, Henry. "The Classification of Duties—Veracity." In *The Methods of Ethics*, bk. 3, chap. 7. Palgrave Macmillan, 1962. https://doi.org/10.1007/978-1-349-81786-3_22.

Squires, Judith. "Public and Private." In *Political Concepts*, edited by Richard Bellamy and Andrew Mason. Manchester University Press, 2018. https://doi.org/10.7765/9781526137562.00015.

"Star-Bulletin Special: The Starr Report." Starbulletin.com. Accessed July 5, 2025. https://archives.starbulletin.com/specials/starr/intro.

Sullivan, Andrew. "Lies That Matter." *New Republic*, 1998. Reproduced at UltraFaction. Accessed July 5, 2025. https://www.ultrafaction.com/articles001/anti-clinton-sullivan-98.html.

U.S. Constitution, amend. I. Constitution Annotated, Library of Congress. https://constitution.congress.gov/constitution/amendment-1/.

Walzer, Michael. "Political Action: The Problem of Dirty Hands." *Philosophy & Public Affairs* 2, no. 2 (1973): 160–80. https://doi.org/10.2307/2265064.

Review & Critique

Yao Ming Resigned and His CBA Reform: Destiny or Inadequacy?

Qunyan Wang
Grade 10

Editors' Introduction

This essay examines Yao Ming's reform of the Chinese Basketball Association and his eventual resignation. It asks whether the outcome reflects the constraints of an entrenched system or the limits of his own leadership. The essay rests on a well-informed account of Chinese basketball culture and uses its source material effectively, which lets it speak with authority about the institutional setting of the reforms. It then turns to ethical theory to evaluate Yao's conduct, weighing the significance of intentions against that of actions and their results.

The essay is anchored in the specific case of Chinese basketball, and the general ethical question it raises is left largely implicit. Its ethical analysis is presented in outline, resting on the distinction between intention and outcome rather than on extended engagement with the literature in moral philosophy.

1. Historical Synopsis

In the early years of the founding of the People's Republic of China, a simple and efficient system was needed to develop the national sports team. It was at that time that the "three-tier athletic development system" was adopted as the national principle, and of course, basketball was a part of the plan. The three tiers, consisting of Sports Traditional Schools (primary and secondary sports schools), Sports Schools (high school level sports schools), and finally, National Training Teams or Provincial Professional Teams, were the designated places for all athletes.[1] The National Basketball League was completely made up of provincial professional teams, and college basketball at that time had no relationship with professional athletes. We can almost say for certain that during that period, the college sports in China were purely "amateur." Imagine this: being born in Akron, Ohio, Lebron James grew up in a "sports school," and could only play for the "Cleveland State Team" (no Cavaliers yet!) in his whole career.

The system did contribute to the rapid development of the national professional sports teams in the beginning. But over time, the system started to hold athletes back. Professional sports and school sports stayed completely separate, which made it harder for talent to develop in a well-rounded way.[2] Athletes had limited chances to grow individually, explore opportunities abroad, or blend sports training with academic education due to the provincial team's full control over them.[3] For example, Yao Ming's journey to the NBA in 2002 was actually very difficult. According to Martin Fackler, Yao was initially forced to hand over half of his income to China's government and sports authorities if he joined the NBA, and

---

[1] Liu Yongfeng et al., "Zhongguo tese lanqiu fazhan zhi lu: Neihan, kunjing yu weilai" 中国特色篮球发展之路：内涵、困境与未来 [Path for basketball development with Chinese characteristics: Connotation, dilemma and future], *Journal of Shanghai University of Sport* 47, no. 7 (2023): 53–63, 86, https://doi.org/10.16099/j.sus.2022.06.01.0007.

[2] Yu Zhenfeng et al., "Zhongguo lanqiu jingji houbei rencai xianzhuang diaocha yu peiyang duice" 中国篮球竞技后备人才现状调查与培养对策 [Study on the status quo and the countermeasure for our reserve basketball talents], *Journal of Physical Education*, no. 5 (2002): 123–26, https://doi.org/10.16237/j.cnki.cn44-1404/g8.2002.05.047.

[3] Fuhua Huang and Fan Hong, "Globalization and the Governance of Chinese Sports: The Case of Professional Basketball," *International Journal of the History of Sport* 32, no. 8 (2015): 1030–43, https://doi.org/10.1080/09523367.2015.1035261.

after tough negotiations, he finally received permission on the day of the draft.[4]

Having lived through the system himself, Yao Ming—one of China's biggest basketball stars—saw both its strengths and its serious flaws. After retiring from the NBA, he took on the challenge of reforming Chinese basketball. In 2017, he was appointed as the President of the CBA (Chinese Basketball Association), and under his leadership, the CBA officially separated from direct government control and set up the CBA Company. Yao's reforms focused on a few big things: making the league more competitive and commercialized (introducing the CBA Draft and promoting player trades), fixing youth training systems ("Little Basketball" programs for kids and developing CUBA (Chinese University Basketball Association)), professionalizing the referee industry, and trying to bring Chinese basketball closer to international standards.[5] His goal was to modify the rigid system while still keeping some of the government support that had helped Chinese sports succeed in the past. However, due to the men's national basketball team's continued decline in performance (for example, failing to qualify for the Tokyo Olympics), Yao Ming has to quit his position in 2024.[6]

2. Ethical Analysis

From the virtue ethics perspective, Yao Ming showed a lot of courage and good judgment in the reform. Growing up under the three-tier system, he knew what needed to be changed and stepped up for real change. He was going against a lot of powerful people and traditions since the players who came from provincial teams were long considered "property" of the team, which was ridiculous in a commercialized league. His clear long-term vision was

---

[4] Martin Fackler, "Beijing Could Claim Half of Chinese Star's NBA Earnings," *Deseret News*, April 25, 2002, archived July 28, 2019, at https://web.archive.org/web/20190728150150/https://www.deseretnews.com/article/910072/Beijing-could-claim-half-of-Chinese-stars-NBA-earnings.html.

[5] Yiyan, "Yao Ming gaige xijie baoguang: Zengjia liansai shijian, xianzhi waiyuan" 姚明改革细节曝光:增加联赛时间限制外援 [Details of Yao Ming's reform revealed: Extended season and restrictions on foreign players], NetEase Sports, February 23, 2017, https://sports.163.com/special/anglezero/yaominggetposition.html.

[6] "Yao Ming ciqu Zhongguo lanqiu xiehui zhuxi zhiwu, Guo Zhenming jieren" 姚明辞去中国篮球协会主席职务 郭振明接任 [Yao Ming resigns as president of the Chinese Basketball Association; Guo Zhenming appointed as successor], *Sina News*, October 31, 2024, https://news.sina.cn/2024-10-31/detail-incumwmy7297999.d.html.

also important during his career as the president. Yao knew that if we wanted the Chinese basketball industry to truly grow, we needed a more supportive environment. That's the reason why he put great effort into developing school basketball programs, pushing for referee professionalization, etc. Lots of people only focus on the current failure of the Chinese men's basketball team and blame it on Yao, but they ignore the fact that reform needs time to see its effect, especially when facing a huge and rigid system that China used to have. "Actions are as important as intentions," and both Yao's actions and intentions are virtuous.[7]

From the deontology point of view, Yao's reforms were the right thing to do because they respected the players' basic rights and did not treat them as tools. Before his changes, a lot of athletes had almost no say over their own careers. Yao fought to fix that by making the CBA more commercialized and giving players clearer chances to move and grow. You can see the impact too: after his reforms, more Chinese players like Zhou Qi and Cui Yongxi got the chance to play in the NBA or in Australia. Players from the women's league, like Li Meng and Han Xu, have been able to join the WNBA and other international leagues and even become star players.[8] We can also see more and more young athletes attending American schools for elite training, and all those things are hard to imagine before his reform.

3. Analytic Conclusion

Let's go back to the topic. What led to the resignation of Yao Ming? I think the "inadequacy" is the lack of time, and the "destiny" is the deep-rooted system. Yao's reforms were ambitious, but structural changes cannot happen in just a few years, especially when he chose the bottom-up approach. But his seven years of effort were far from useless. More and more players start to get benefits from competing and training abroad, and more players are pursuing opportunities in the NBA Draft, such as Zeng FanBo and Cui Yongxi. Athletes' performances from CUBA are becoming more and more competitive, and the first pick of

---

[7] Brian Boone, *Ethics 101: From Altruism and Utilitarianism to Bioethics and Political Ethics, an Exploration of the Concepts of Right and Wrong* (Adams Media, 2017), chap. 1, "The Moral Philosophy of Aristotle."

[8] Xu Bangyin, "Li Meng denglu WNBA, Zhongguo nülan yi you 4 wei xianyi guoshou jiameng" 李梦登陆 WNBA，中国女篮已有 4 位现役国手加盟 [Li Meng joins the WNBA, bringing the number of active Chinese women's national team players in the league to four], *Xin jing bao* [Beijing News], March 14, 2023, reposted on Sohu, https://www.sohu.com/a/653981949_114988.

2023, Chen Guohao, exploded for 20 points, 10 rebounds, and 5 blocks in his CBA debut. The national female basketball team has already risen to second place in the world, and people are paying more attention and money to the WCBA (Women's Chinese Basketball Association).[9] Young athletes are entering D1 NCAA schools, such as Dan Houran and Wang Junjie.[10] The seeds have already been planted, and all they need is time to blossom.

On the other hand, Yao's reforms harmed the interests of many traditional strong teams that held onto their star players using the regulations. The clash between CBA star Zhou Qi and the Xinjiang Guanghui Basketball Club is an example. In 2022, Zhou Qi and the Xinjiang Guanghui Basketball Club had a major dispute over the contract. Zhou Qi wanted more flexibility to pursue opportunities overseas, but the club refused to let Zhou Qi go. The disagreement escalated when Xinjiang refused to release Zhou Qi or offer a fair contract, leading to his decision to sit out the CBA season. In 2023, with tensions still running high over salary cap rules and Yao Ming's push for player rights reforms, the Xinjiang team suddenly pulled out of the CBA, claiming "CBA's unfair behavior."[11] Many people, particularly local fans from Xinjiang, viewed this incident as another piece of proof of the difficulties—and, in their eyes, the failure—of Yao Ming's reform efforts.

The main idea of Yao Ming's reforms was a simple but important idea: athletes deserve the right to control their own careers. The traditional system has treated players like property, tying them to teams and limiting their choices for too long. Yao tried to make the CBA a place where athletes could have real agency—the right to move, to compete, and to grow without being trapped within certain teams. This matters because rights are the foundation of any healthy sports system, and without them, real development is impossible. Players' hard work should lead them to better opportunities, not just for the benefit of a few teams. Yao's resignation shows how hard it is to change a system that resists giving up control, but the changes he made—more chances abroad and more respect for athletes—are

---

[9] Xu, "Li Meng denglu WNBA."

[10] Jin de jiu shi ni, "Zhongguo nanlan liuyang qiuyuan baogao (er): Liang ren huode NCAA shoufa jihui" 中国男篮留洋球员报告（二）：两人获得 NCAA 首发机会 [Report on Chinese basketball players abroad (part II): Two players earn starting positions in the NCAA], Sohu, December 25, 2024, https://m.sohu.com/a/841520765_122028124/?pvid=000115_3w_a.

[11] Wang Yayun, "Xinjiang nanlan xuanbu tuichu CBA, shengcheng wei Zhou Qi huafei jin 5000 wan" 新疆男篮宣布退出 cba，声称为周琦花费近 5000 万 [Xinjiang basketball team announces withdrawal from the CBA, claims to have spent nearly 50 million yuan on Zhou Qi], *Nanfang dushi bao* [Southern Metropolis Daily], March 1, 2023, https://m.mp.oeeee.com/a/BAAFRD000020230301768867.html.

not going away. His work matters because it opened a door that cannot easily be closed again. We can see more and more athletes seeking opportunities both through transfers and playing abroad. After all, fighting for athletes' rights is about fairness, dignity, and the future of Chinese basketball.

Bibliography

Boone, Brian. *Ethics 101: From Altruism and Utilitarianism to Bioethics and Political Ethics, an Exploration of the Concepts of Right and Wrong*. Adams Media, 2017.

Fackler, Martin. "Beijing Could Claim Half of Chinese Star's NBA Earnings." *Deseret News*, April 25, 2002. Archived July 28, 2019, at https://web.archive.org/web/20190728150150/https://www.deseretnews.com/article/910072/Beijing-could-claim-half-of-Chinese-stars-NBA-earnings.html.

Huang, Fuhua, and Fan Hong. "Globalization and the Governance of Chinese Sports: The Case of Professional Basketball." *International Journal of the History of Sport* 32, no. 8 (2015): 1030–43. https://doi.org/10.1080/09523367.2015.1035261.

Jin de jiu shi ni. "Zhongguo nanlan liuyang qiuyuan baogao (er): Liang ren huode NCAA shoufa jihui" 中国男篮留洋球员报告（二）：两人获得 NCAA 首发机会 [Report on Chinese basketball players abroad (part II): Two players earn starting positions in the NCAA]. Sohu, December 25, 2024. https://m.sohu.com/a/841520765_122028124/?pvid=000115_3w_a.

Liu Yongfeng, Chen Rui, Zhang Mingxin, et al. "Zhongguo tese lanqiu fazhan zhi lu: Neihan, kunjing yu weilai" 中国特色篮球发展之路：内涵、困境与未来 [Path for basketball development with Chinese characteristics: Connotation, dilemma and future]. *Journal of Shanghai University of Sport* 47, no. 7 (2023): 53–63, 86. https://doi.org/10.16099/j.sus.2022.06.01.0007.

Sina News. "Yao Ming ciqu Zhongguo lanqiu xiehui zhuxi zhiwu, Guo Zhenming jieren" 姚明辞去中国篮球协会主席职务 郭振明接任 [Yao Ming resigns as president of the Chinese Basketball Association; Guo Zhenming appointed as successor]. October 31, 2024. https://news.sina.cn/2024-10-31/detail-incumwmy7297999.d.html.

Wang Yayun. "Xinjiang nanlan xuanbu tuichu CBA, shengcheng wei Zhou Qi huafei jin 5000 wan" 新疆男篮宣布退出 cba，声称为周琦花费近 5000 万 [Xinjiang basketball team announces withdrawal from the CBA, claims to have spent nearly 50 million yuan on Zhou Qi]. *Nanfang dushi bao* [Southern Metropolis Daily], March 1, 2023. https://m.mp.oeeee.com/a/BAAFRD000020230301768867.html.

Xu Bangyin. "Li Meng denglu WNBA, Zhongguo nülan yi you 4 wei xianyi guoshou jiameng" 李梦登陆 WNBA，中国女篮已有 4 位现役国手加盟 [Li Meng joins the

WNBA, bringing the number of active Chinese women's national team players in the league to four]. *Xin jing bao* [Beijing News], March 14, 2023. Reposted on Sohu. https://www.sohu.com/a/653981949_114988.

Yiyan. "Yao Ming gaige xijie baoguang: Zengjia liansai shijian, xianzhi waiyuan" 姚明改革细节曝光:增加联赛时间限制外援 [Details of Yao Ming's reform revealed: Extended season and restrictions on foreign players]. NetEase Sports, February 23, 2017. https://sports.163.com/special/anglezero/yaominggetposition.html.

Yu Zhenfeng, Zhang Zhendong, Zhang Jianjun, et al. "Zhongguo lanqiu jingji houbei rencai xianzhuang diaocha yu peiyang duice" 中国篮球竞技后备人才现状调查与培养对策 [Study on the status quo and the countermeasure for our reserve basketball talents]. *Journal of Physical Education*, no. 5 (2002): 123–26. https://doi.org/10.16237/j.cnki.cn44-1404/g8.2002.05.047.

Review & Critique

Beyond Russell’s Formula: A Critical Reassessment of Rousseau and Locke’s Influence on Hitler, Roosevelt, and Churchill

Minghao Liu
Grade 11

Editors' Introduction

This essay revisits Bertrand Russell's remark in *A History of Western Philosophy* that Hitler is an outcome of Rousseau, while Roosevelt and Churchill are outcomes of Locke. It compares the historical record of each leader with the political philosophy he is said to inherit, and it gives a detailed account of where Russell's pairing holds and where it fails. On this basis, the essay offers a creative critique of Russell's historiography. It argues that Russell uses the opposition between these twentieth-century leaders to stage an incompatibility between Rousseau and Locke, and that this approach is ultimately limited as a way of writing the history of ideas.

The essay devotes most of its length to comparing each leader with his philosophical counterpart, and its central historiographical thesis comes fully into view in the closing paragraphs. These comparisons rest more on the historical record than on close reading of Rousseau's and Locke's own texts.

In *A History of Western Philosophy*, Bertrand Russell famously claimed that "Hitler is an outcome of Rousseau; Roosevelt and Churchill of Locke."[1] At first glance, this statement assigns historical responsibility to two Enlightenment philosophers for shaping the moral and political landscapes of the 20th century. Yet such a linkage demands closer scrutiny. Hitler was less a direct outcome of Rousseau than a misinterpretation of Nietzsche. Roosevelt and Churchill share a similar ideology that mostly coincides with Locke's philosophy. Roosevelt made adjustments to Lockean ideals by practising positive liberties, and Churchill engaged in imperialism that goes against Lockean principles but aligns with Locke's personal choices. Meanwhile, Russell's biases as a result of his context also influenced his limited interpretation.

We will first examine the quote itself, analysing the misattribution of Rousseau to Hitler, Roosevelt's extension of Locke's principles, and Churchill's almost verbatim adherence to Locke, before moving on to investigate Russell's context and how it influenced his stance.

Russell's attribution of Rousseau's philosophy to Hitler was a generalisation and misrepresentation. The most obvious contention is Rousseau's idea of the General Will and that individuals should adhere to a collective will aimed at the common good.[2] Freedom was obedience to a law one has prescribed for oneself. Individuals must be compelled to follow the General Will for freedom to endure. An ideal society would be one where people were "forced to be free."[3] Individuals can only discern the General Will through direct deliberation, which is lacking in representative governments. Rousseau rejects representative governments in favour of direct participation and enacting policy through elective aristocracy: the Sovereignty of the People.[4]

Drawing parallels between Rousseau's political philosophy and Hitler's actions can be tempting. Hitler ascended to power through the votes of the German people. The emergence of the National Socialist Party as the largest party allowed Hitler to become Chancellor; this was a direct result of the democratic principles of the Weimar Republic. Following his appointment, Hitler merged the offices of Chancellor and President, becoming

---

[1] Bertrand Russell, *A History of Western Philosophy* (Simon and Schuster, 2008), 685.

[2] Jean-Jacques Rousseau, *A Treatise on the Social Compact; or, The Principles of Political Law* (London: D. I. Eaton, 1795), bk. 1, chap. 6; bk. 2, chap. 3.

[3] Rousseau, *Treatise on the Social Compact*, bk. 1, chap. 7.

[4] Rousseau, *Treatise on the Social Compact*, bk. 3, chaps. 5, 15.

de facto dictator.[5] One could argue that German voters, by electing Hitler into power, imposed upon themselves a General Will soon to be enforced by the aristocratic elite of Hitler and his inner circle. Therefore, they were indeed "forced" to be free after Hitler consolidated power and restricted civil liberties.

However, the alignment between Hitler's actions and Rousseau's philosophy is minimal. Hitler was a pragmatist, acting in the most advantageous way to consolidate power. He had always made clear, from his days in a Bavarian prison writing *Mein Kampf*[6] to his final moments in the Berlin bunker,[7] that he held a disdain for democracy, instead placing his faith in the Führerprinzip—a political hierarchy where absolute power resides in one strong leader. During Hitler's rise to power, democratic principles were merely a means to an end, an end that led to the demise of those very principles.

The more jarring difference between Rousseau and Hitler lies in the racial policies of the National Socialist movement, which were equally, if not more important than, their political ideology. Rousseau had never indicated that racial superiority was a commendable policy.[8]

The same contradictions exist regarding Hitler's economic policies. While Rousseau criticised liberal capitalism and attributed inequality to property ownership and class division,[9] this view circulated among socialists and communists, whom Hitler opposed.[10] Nazi Germany's economic structure was more akin to a statist, autarkic corporatist or oligarchic state existing primarily for the waging of war, with labour unions outlawed and industrial elites such as IG-Farben and Krupp becoming major political influencers.[11]

The next question readily presents itself: Who, then, if not Rousseau, influenced Hitler? If we must identify one philosopher, it would be Nietzsche. Nietzsche believed in the Übermensch ("Superman")—someone who creates their values and transcends conventional morality.[12] Central to his belief was the idea of the Will to Power, a drive in all life to grow

---

[5] William L. Shirer, *The Rise and Fall of the Third Reich: A History of Nazi Germany* (Simon and Schuster, 1990).

[6] Adolf Hitler, *Mein Kampf*, trans. Ralph Manheim (1925; repr., Pimlico, 1992).

[7] Shirer, *Rise and Fall of the Third Reich*.

[8] Rousseau, *Treatise on the Social Compact*.

[9] Jean-Jacques Rousseau, *A Discourse on Inequality*, trans. Maurice Cranston (Penguin, 1985).

[10] Shirer, *Rise and Fall of the Third Reich*.

[11] Shirer, *Rise and Fall of the Third Reich*.

[12] Friedrich Nietzsche, *Thus Spoke Zarathustra* (CreateSpace Independent Publishing Platform, 2018).

and overcome challenges.[13] Hitler appropriated these ideas to lend his regime a veneer of legitimacy, visiting Nietzsche's archives in 1934 and maintaining a close relationship with Nietzsche's sister Elisabeth.[14] Gradually, Hitler transformed the Übermensch into a racial idol and the Will to Power into a justification for authoritarianism. The National Socialists' selective usage of Nietzsche's ideas ignored his suspicion of the state and condemnation of antisemitism to exploit the name of a respected German philosopher. Hitler was more a pragmatist than someone who followed a certain philosophical dogma. His worldview was built from his experiences on the fronts of World War I, mingling at political meetings, and the "bitter loser" sentiments shared by most Germans at the time, representing a patchwork of Social Darwinism, Pan-German nationalism, and antisemitism rather than the coherent philosophy of any particular thinker.

On the other hand, Locke's association with Roosevelt is more straightforward. Most of his political ideology and policies were either the result of or an extension of Locke's arguments. Locke believed that individuals were born with the natural rights of life, liberty, and property, and that it was the government's duty to protect these rights.[15] He was also a proponent of the social contract, where individuals consent to form a government by relinquishing certain freedoms for the protection of their rights. Roosevelt served as a defender of democracy and popular participation in a polarised world troubled by fascism and communism.

Roosevelt was raised and entered a political career influenced by Lockean philosophy and classical liberalism. In 1931, as governor of New York, he persuaded the Republican-controlled legislature to establish the Temporary Emergency Relief Administration, which offered unemployment benefits to 10 per cent of New York families.[16] All these actions reflect his unwavering belief in the Social Contract and the obligations public officials have to improve the conditions of their constituents. It is also important to note his other influences: during his time at Harvard, he was inspired by his cousin, Theodore Roosevelt, who had advocated for progressive reforms.[17] He was also influenced by his wife, Eleanor Roosevelt,

---

[13] Friedrich Nietzsche, *The Will to Power: An Attempted Transvaluation of All Values*, trans. Anthony M. Ludovici (T. N. Foulis, 1913).

[14] Shirer, *Rise and Fall of the Third Reich*.

[15] John Locke, *Two Treatises of Government* (Standard Ebooks, 2025).

[16] Jean Edward Smith, *FDR* (Random House Trade Paperbacks, 2008).

[17] James MacGregor Burns, *Roosevelt: The Soldier of Freedom* (Smithmark, 1996).

who engaged in charitable work for the impoverished in New York and opened her husband's eyes to the dire conditions in New York slums.[18]

It is important to recognise that while Locke advocated Negative Liberty (freedom *from* interference by others), Roosevelt believed in Positive Liberty (freedom *to* pursue a satisfactory livelihood). Between when he assumed office in 1933 until his death in 1945, Roosevelt made several adjustments to Locke's notion of natural rights and advocacy for a minimal government role by adopting a more proactive approach in safeguarding natural rights, creating economic and social conditions that enabled genuine freedom, and granting privileges such as the right to work, housing, and security. Acting from experience and learning from the failures of the non-intervention policies set out by Herbert Hoover, this was a groundbreaking move that managed to save the nation. It can be easy to see disparities between Roosevelt's policies and what Locke advocated, as a few were deemed unconstitutional. These included the National Industrial Recovery Act, which granted the executive branch excessive legislative power, and the Agricultural Adjustment Act, which breached the Tenth Amendment.[19] Roosevelt overstepped in several areas and extended federal reach to a point that was controversial in the face of classical Lockean thought. Yet the second round of New Deal policies saw less backlash, indicating that most were aware of the benefits of extending Lockean thought. Roosevelt's policies could be more accurately described as an evolution of Locke's principles, signifying a shift from classical liberalism to modern liberalism to fit the circumstances of the times.

Like Roosevelt, Churchill was also a defender of Lockean ideals. Though Churchill was brought up following the Anglican faith, he expressed in a letter that he preferred Protestantism over Roman Catholicism because he felt it "a step nearer Reason"[20]—aligning himself both with the Protestant beliefs and trust in reason of Locke. Churchill has been described by his biographers as a fundamental conservative with a liberal outlook.[21] Despite frequent changes in party affiliation during his political career, he was fully subscribed to Locke's ideas of a limited government. He was staunchly anti-socialist due to the movement's advocacy for government intervention in the market, and opposed the

---

[18] Allan M. Winkler, *Franklin D. Roosevelt and the Making of Modern America* (Pearson Longman, 2006).

[19] Smith, *FDR*.

[20] Martin Gilbert, *Churchill: A Life* (Rosetta Books, 2014).

[21] Gilbert, *Churchill*.

conservatives' promotion of protectionism, leading to his first defection to the Liberal Party in 1904. He was a firm believer in the English parliamentary democracy stemming from the Glorious Revolution, as he led his nation throughout the war against fascism.

Unfortunately, unlike Roosevelt, Churchill also inherited many of Locke's mistakes, especially concerning colonialism and imperialism. Locke had invested in the Royal African Company,[22] which traded enslaved Africans, and helped draft the 1669 Constitution of Carolina,[23] which permitted hereditary nobility and slavery. Similar to Locke, Churchill preached principles of freedom and equality, but at the same time was also an ardent imperialist in practice. Churchill's response to the Bengal famine was criticised by contemporaries as insufficient. In his private remarks made to Secretary for India Leo Amery, Churchill concluded aid would be pointless because "Indians [were] breeding like rabbits."[24] Churchill saw British Imperialism as a form of altruism, viewing the devolution of powers to Commonwealth members as irrational. This is manifested in his opposition to the Government of India Act of 1935 and his dislike of Mahatma Gandhi.[25] While these views may seem to contradict Lockean values of freedom and self-determination, it is more apt to say that Churchill inherited Locke's ideals but also shared in his blind spots. Both failed to consistently uphold their beliefs, instead reverting to a more pragmatic approach as the situation suited them.

Having established many inconsistencies between the philosophers and the political leaders, we must question Russel's credibility and motivations while writing this.

Russell appears to place significant emphasis on the supposed incompatibility between the democratic philosophies of Rousseau and Locke. In *A History of Western Philosophy*, he opens his chapter on Rousseau by drawing a contrast between the two thinkers, concluding that they "saw no compatibility." Immediately following this statement, he introduces the claim mentioned in the prompt.[26] Indeed, Locke and Rousseau diverged

---

[22] Holly Brewer, "Slavery, Sovereignty, and 'Inheritable Blood': Reconsidering John Locke and the Origins of American Slavery," *American Historical Review* 122, no. 4 (2017): 1038–78, https://doi.org/10.1093/ahr/122.4.1038.

[23] James Farr, "'So Vile and Miserable an Estate': The Problem of Slavery in Locke's Political Thought," *Political Theory* 14, no. 2 (1986): 263–89, https://doi.org/10.1177/0090591786014002005.

[24] Yogita Limaye, "Churchill's Legacy Leaves Indians Questioning His Hero Status," BBC News, July 21, 2020, https://www.bbc.com/news/world-asia-india-53405121.

[25] Gilbert, *Churchill*.

[26] Russell, *History of Western Philosophy*, 685.

fundamentally in their ideas of democracy. Yet, while the ideologies of Hitler on one hand and Roosevelt and Churchill on the other were starkly opposed—much like those of Rousseau and Locke—it would be a logical error to infer a causal relationship based on this contrast alone. The observant reader also notices that Russell wrote the book during World War II and published it soon after it ended. Thus, Russell simply did not possess the historical resources to formulate such a conclusion. More importantly, he lacked the benefit of hindsight, with many historians considering 10-20 years a required time to elapse between the time of the event and the time of analysis for accurate historical research.[27] Much of the material concerning Nazi Germany was scattered around the world and took decades until historians could analyze them. Many government documents in the US and the UK were not declassified until years later. Much of the commentaries of the time, including those by Hayek,[28] Druker,[29] and Nisbet,[30] shared similar opinions to Russell, also working with limited historical resources. In contrast, much of the work published several decades after World War II, such as that by Birzer[31] and Wokler,[32] offered opposing views, especially concerning Russell's interpretation of Rousseau's influence on Hitler. In this light, Russell's sweeping assertion may be better seen as a rhetorical flourish rather than a rigorously argued position grounded in historical or philosophical analysis.

Ultimately, Russell's assertion is more of a rhetorical provocation influenced by the historical context in which he wrote than a precise genealogical assertion. The statement is best understood as an invitation to investigate how ideas can be interpreted or distorted by context rather than as a definitive mapping of Enlightenment thought onto leadership in the 20th century. History rarely follows neat philosophical lines; rather, it reflects the messiness of human actions and the distortions of time. Reading Russell's assertion as history runs the risk of being oversimplified; reading it critically marks the start of the actual historical process.

---

[27] Marc Bloch, *The Historian's Craft*, trans. Peter Putnam (Manchester University Press, 1954).

[28] F. A. Hayek, *The Road to Serfdom: Text and Documents; The Definitive Edition* (Routledge, 2014).

[29] Peter F. Drucker, *The Future of Industrial Man* (Transaction Publishers, 1995).

[30] Robert A. Nisbet, "Rousseau and Totalitarianism," *Journal of Politics* 5, no. 2 (1943): 93–114, https://doi.org/10.2307/2125667.

[31] Bradley J. Birzer, "Rousseau's Collectivism," *The Imaginative Conservative*, January 4, 2020, https://theimaginativeconservative.org/2020/01/jean-jacques-rousseau-collectivism-bradley-birzer.html.

[32] Robert Wokler, *Rousseau: A Very Short Introduction* (Oxford University Press, 2001).

Bibliography

Birzer, Bradley J. "Rousseau's Collectivism." *The Imaginative Conservative*, January 4, 2020. https://theimaginativeconservative.org/2020/01/jean-jacques-rousseau-collectivism-bradley-birzer.html.

Bloch, Marc. *The Historian's Craft*. Translated by Peter Putnam. Manchester University Press, 1954.

Brewer, Holly. "Slavery, Sovereignty, and 'Inheritable Blood': Reconsidering John Locke and the Origins of American Slavery." *American Historical Review* 122, no. 4 (2017): 1038–78. https://doi.org/10.1093/ahr/122.4.1038.

Burns, James M. *Roosevelt: The Soldier of Freedom*. 1970. Reprint, Smithmark, 1996.

Drucker, Peter F. *The Future of Industrial Man*. 1942. Reprint, Transaction Publishers, 1995.

Farr, James. "'So Vile and Miserable an Estate': The Problem of Slavery in Locke's Political Thought." *Political Theory* 14, no. 2 (1986): 263–89. https://doi.org/10.1177/0090591786014002005.

Gilbert, Martin. *Churchill: A Life*. 1991. Reprint, Rosetta Books, 2014.

Hayek, Friedrich A. *The Road to Serfdom*. 1944. Reprint, Routledge, 2014.

Hitler, Adolf. *Mein Kampf*. Translated by Ralph Manheim. 1925. Reprint, Pimlico, 1992. https://greatwar.nl/books/meinkampf/meinkampf.pdf.

Limaye, Yogita. "Churchill's Legacy Leaves Indians Questioning His Hero Status." BBC News, July 21, 2020. https://www.bbc.com/news/world-asia-india-53405121.

Locke, John. *Two Treatises of Government*. 1689. Reprint, Standard Ebooks, 2025.

Nietzsche, Friedrich. *Thus Spoke Zarathustra*. 1883. Reprint, CreateSpace Independent Publishing Platform, 2018.

Nietzsche, Friedrich. *The Will to Power: An Attempted Transvaluation of All Values*. Translated by Anthony M. Ludovici. 1901. Reprint, T. N. Foulis, 1913.

Nisbet, Robert A. "Rousseau and Totalitarianism." *Journal of Politics* 5, no. 2 (1943): 93–114. https://doi.org/10.2307/2125667.

Rousseau, Jean-Jacques. *A Discourse on Inequality*. Translated by Maurice Cranston. 1755. Reprint, Penguin Books, 1985.

Rousseau, Jean-Jacques. *A Treatise on the Social Compact; or, The Principles of Political Law*. London: D. I. Eaton, 1795.

Russell, Bertrand. *A History of Western Philosophy*. Simon and Schuster, 2008.

Shirer, William L. *The Rise and Fall of the Third Reich: A History of Nazi Germany*. 1960. Reprint, Simon and Schuster, 1990.

Smith, Jean E. *FDR*. Random House Trade Paperbacks, 2008.

Winkler, Allan M. *Franklin D. Roosevelt and the Making of Modern America*. Pearson Longman, 2006.

Wokler, Robert. *Rousseau: A Very Short Introduction*. Oxford University Press, 2001.

Research Paper

Semantic Stability and Floating Meaning: Thinking About the Language of AI

Shenzhi Zhang
Grade 11

Editors' Introduction

This paper revisits Wittgenstein's claim that meaning is use and argues that large language models, because they produce language apart from human practice, generate text whose form stays stable while its meaning can "float" across contexts. To describe this condition, it proposes two original tools. The Prototypical Semantic Core is the minimal cluster of features that keeps a word recognizable across uses. Semantic Attachment Levels is a four-level scale of how firmly a use is anchored in practice and in wider conceptual networks. On this basis, the paper describes LLMs as semantic catalysts rather than independent creators of concepts. It closes with a practical proposal, "adversarial inquiry," in which users treat a model's output as a claim to be tested rather than accepted.

The paper sets out its two tools mainly through definition, and concrete illustrations appear later in the argument. Its discussion of the cognitive risks of relying on LLMs draws on recent empirical studies whose findings are still preliminary. Alternative accounts of human–machine cognition, such as the extended mind, enter mainly in the conclusion; the framework is not tested against them as rival views.

1. Introduction

The philosophical question of whether machines can truly understand language has a long history. Douglas Hofstadter, in *I Am a Strange Loop*, depicts a hypothetical dialogue between Socrates and Plato that probes this issue. In the dialogue, Plato attempts to distinguish life and thought from a machine capable of flawlessly answering questions, arguing that such a machine merely manipulates symbols mechanically without genuine understanding. Socrates challenges this distinction by suggesting that much of human language use may also be automatic or habitual, thereby questioning whether the boundary between human consciousness and machine processing is as clear as it first appears.[1]

This hypothetical dialogue captures the enduring puzzle of understanding and meaning. A related and influential thought experiment is John Searle's Chinese Room, which imagines a person inside a room following syntactic rules to manipulate Chinese characters without understanding their meaning.[2] The experiment challenges the assumption that successful linguistic performance necessarily implies semantic understanding. More broadly, it highlights a possible gap between manipulating signs and understanding what those signs signify.

Large Language Models complicate rather than clarify this old philosophical problem. They produce fluent sentences, yet critics continue to question whether such systems have access to practice-grounded meaning. As Bender and Koller argue, "a system trained only on form has no way to learn meaning" in the full communicative sense.[3] The key challenge, then, is not simply whether AI fails to "mean," but how stable its meanings remain when language moves across different contexts. Here I ask three more specific questions: Can AI maintain semantic stability across different contexts of use? How should we describe different degrees of semantic stability? And how should we treat the language generated by AI?

My contention is that AI can reproduce linguistic rules and structures while meaning becomes less stable when the context that normally constrains those structures is absent or

---

[1] Douglas R. Hofstadter, *I Am a Strange Loop* (Basic Books, 2007).

[2] John R. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–57.

[3] Emily M. Bender and Alexander Koller, "Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (Association for Computational Linguistics, 2020), 5185.

changed. I call this condition *floating meaning*. Floating meaning does not mean that AI-generated language is completely meaningless. An LLM-generated sentence can become meaningful when human speakers interpret and use it within an established practice. The problem is instead that the form of a word and its broad semantic associations can travel more easily than the specific rules governing its use.

I offer two analytical tools to illustrate this problem: PSC and SAL. These tools allow us to examine where semantic continuity remains and where attachment to context becomes weak. My central claim is that current AI is neither an autonomous maker of concepts nor an understanding subject in the same sense as a human participant in linguistic practice. It is better understood as a catalyst that can support human inquiry by generating possible conceptual relations that human users must then examine and stabilize.

This argument is built on several premises drawn from the philosophy of language and the philosophy of science. The starting point is Wittgenstein's insight that meaning is closely tied to use. Shared rules within human practices stabilize a word's sense, while outside these practices' symbols can become less determinate.[4] This does not imply that every word has only one correct meaning. Rather, it means that the stability of a meaning depends on the ways in which speakers learn to apply, distinguish, correct, and respond to uses of the word.

This distinction is important for LLMs. Human practices produce language: scientific practice produces scientific writing, legal practice produces legal texts, and social interaction produces ordinary discourse. LLMs are trained mainly on the linguistic products of these practices. They therefore acquire many genuine semantic relations, but they encounter them primarily through patterns represented in language rather than through direct participation in every practice that produced those patterns. The difference is not simply between "meaningful humans" and "meaningless machines." It is between learning patterns produced by a practice and participating in the full set of rules and consequences that stabilize the practice.

This helps clarify the disconnect between form and meaning identified by Bender and Koller. Linguistic form is not empty: texts already contain traces of human conceptual practices. Yet the same form can appear in different fields under different rules. When a word moves between contexts, some of its semantic features may travel successfully while others

[4] Ludwig Wittgenstein, *Philosophical Investigations*, rev. 4th ed., trans. G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte (Wiley-Blackwell, 2009), §43.

should not. Semantic drift occurs when the transferable part survives but the contextual restrictions weaken.

At the same time, this disconnect creates the possibility of conceptual novelty. New concepts are often not completely disconnected from old ones. They emerge when existing ideas are reorganized within new problems and new fields of practice. Kuhn's account of scientific revolutions suggests that genuinely important conceptual changes occur together with changes in problems, methods, and standards of inquiry.[5] A new term becomes a working concept only when a community learns how to use it, what problems it addresses, and how its application can be criticized. Lakoff similarly emphasizes that concepts are not held in isolation but sustained through wider networks of related categories and cultural frames.[6]

These premises explain why LLMs can maintain syntactic coherence and a considerable amount of semantic structure while still risking semantic instability. They can reproduce and recombine relations expressed in human language, but they do not independently establish the practices that determine how a new combination should be used. For this reason, they function more plausibly as catalysts for inquiry than as autonomous creators of meaning.

Current LLMs are generally trained through next-token prediction over very large corpora using Transformer architectures.[7] Scaling data, computation, and model parameters has further improved fluency and pattern recognition.[8] This training allows models to capture complex relationships among words and concepts, but it does not itself place the model inside the practices that originally stabilized those relationships. Post-training can improve reasoning and contextual performance, but the basic distinction remains: reproducing linguistic structure is not identical to participating in all the practical conditions under which that structure acquires determinate use.

Hence, linguistic form can travel more easily than context-specific meaning, and semantic drift appears when tasks or contexts change. To analyze this instability with greater precision, I propose two tools.

---

[5] Thomas S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, 1962).

[6] George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind* (University of Chicago Press, 1987).

[7] Ashish Vaswani et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems 30* (2017).

[8] Jared Kaplan et al., "Scaling Laws for Neural Language Models," preprint, arXiv:2001.08361, 2020.

2. Tool A: Prototypical Semantic Core (PSC)

The *Prototypical Semantic Core* is a small, cross-context anchor for a word. It consists of central semantic features that remain recognizable across different uses. However, PSC should not be treated as a complete definition or a fixed essence. Such an account would conflict with Wittgenstein's own discussion of family resemblance, in which different uses of a concept may overlap without sharing one necessary feature.[9] PSC is therefore better understood as a prototype: a flexible cluster of recurring semantic associations.

The PSC for "transparency," for example, may include access to information, visibility, openness, or the possibility of inspection. These features allow the word to travel across corporate ethics, politics, computing, and other fields while remaining recognizable. At the same time, the PSC limits what the word can plausibly signify. If a use loses all relation to access, visibility, or inspection, readers will begin to question whether "transparency" remains an appropriate term.

PSC therefore provides partial semantic stability. It explains why a word can move across contexts without becoming completely different each time. Yet it does not determine how the word should function in each particular field. Transparency in corporate governance may refer to disclosure and accountability; in cryptography, it may refer to auditability and verifiability while still allowing private information to remain protected. The PSC survives, but the rules of use change.

This limitation is especially important for understanding LLMs. A model may remain close to the PSC of a term and still use that term incorrectly. In such cases, the output sounds plausible precisely because the general semantic associations remain intact. What has weakened is not the entire meaning of the word but its attachment to a particular practice.

While PSC identifies the relatively stable and portable part of meaning, a second tool is needed to describe how securely a particular use is anchored in context.

3. Tool B: Semantic Attachment Levels (SAL)

Semantic Attachment Levels indicate how tightly a particular use of language is attached to practice. SAL is not a test of truth, nor is it a mathematical measure. Instead, it describes increasing degrees of practical and conceptual constraint.

At the lowest level, *L0*, lies surface form. A word fits grammatically into a sentence

[9] Wittgenstein, *Philosophical Investigations*, §§65–71.

and may resemble established discourse, but it is not clearly attached to a specific practice. The language is formally coherent without establishing what rules determine its application.

At *L1*, the use stays close to a shared PSC. The reader recognizes the general meaning of the term, but its specific conditions of application remain unclear. Semantic drift is reduced because the word retains a recognizable core, but several interpretations may still remain possible.

Attachment deepens at *L2*, where a term is tied to a concrete field of tasks, rules, and consequences. Its meaning becomes constrained by what participants in that field actually do. The concept is no longer sustained mainly by general associations; it is connected to recognizable practices.

At *L3*, the term is embedded within a wider network of neighboring concepts, institutions, and systems of correction. A concept at this level is constrained not only by what people do but also by its relation to other concepts and by rules that distinguish correct from incorrect use. Meaning becomes more resilient because changing the use of one term affects its place within a larger system.

SAL does not claim that an L3 use is necessarily true, nor that its meaning can never change. Scientific, political, and philosophical concepts can remain contested even when they are deeply embedded in practice. SAL describes the strength of attachment rather than permanent semantic certainty.

PSC and SAL therefore serve different functions. PSC explains how a word retains enough continuity to travel. SAL describes how securely a particular use is attached after the word enters a context. Semantic drift appears when the form and PSC remain recognizable but the attachment to context weakens.

Consider the word “transparency.” In corporate ethics, it may refer to disclosure rules, audit trails, and accountability to investors and the public. In cryptography, it may point to auditability and verifiability: algorithms may be open to inspection while keys remain protected. In politics, it can either function as a broad ideal of openness or become attached to concrete institutional mechanisms such as disclosure procedures. The word remains recognizably related across these contexts, but its practical implications differ.

LLMs may fuse “publicity” with “proof” or “disclosure” with “verification.” That is a form of semantic drift: the linguistic sign and part of its semantic core remain, but the relations governing the concept shift across contexts. More precisely, this can be called *contextual misattachment*. The problem is not that the word suddenly has no meaning; it is that the model has attached a recognizable meaning to the wrong conceptual network.

This distinction is also important when discussing reasoning. LLMs are optimized during pretraining for statistical prediction rather than directly for logical validity. Their training enables them to learn patterns associated with reasoning, and contemporary models can perform substantial reasoning tasks. However, predictive success does not guarantee logical consistency. A model may produce an internally fluent argument while failing to preserve the exact inferential rules required by a particular context. Logical error and semantic drift are therefore related but not identical. A sentence may be semantically stable but logically invalid, just as a sentence may be logically coherent while employing a concept in the wrong context.

For this reason, it would be too strong to conclude that LLMs cannot reason or cannot carry meaning. The more limited conclusion is that they do not independently guarantee either logical consistency or semantic attachment. Their outputs should therefore be treated as provisional and open to further examination.

4. LLMs as Semantic Catalysts

Given this instability, it is more useful to treat LLMs as catalytic tools. There are two main ways in which they can help us reconsider how meaning and identity are constructed.

First, in generating new writing, LLMs often bring together words and concepts from different fields. Because they are trained on extremely diverse bodies of language, they can identify patterns of association that may not be immediately visible to an individual user. These connections do not prove that the concepts genuinely belong together, but they can suggest new directions for inquiry and force human users to consider unfamiliar conceptual combinations.

However, such combinations initially remain semantically weak. They may be fluent in form without yet having a clear role in practice. A novel relation between two concepts becomes philosophically useful only when a human user asks what the relation means, where it applies, and whether it survives criticism. The catalytic moment therefore requires a human response.

The user does not passively receive the output but reconstructs it through situated knowledge, critical judgment, and practical intention. This process follows directly from PSC and SAL. The model may provide a conceptual combination that retains recognizable semantic cores, but the human user must determine whether the combination can move from weaker to stronger levels of attachment. The model supplies a possible connection; practice

determines whether the connection becomes stable.

This also explains why a novel combination should not immediately be called a new concept. If meaning depends on use, then a concept becomes established only when users can determine its field of application, distinguish appropriate from inappropriate uses, connect it to neighboring concepts, and revise it when necessary. Kuhn's account of conceptual change supports this point: new scientific concepts become meaningful through their role in new practices of inquiry, not through linguistic invention alone.

LLMs can therefore participate in conceptual creation without independently completing it. They can propose combinations, formulations, or distinctions. But human practices must stabilize these possibilities. For this reason, LLMs function fundamentally as *semantic catalysts rather than autonomous concept creators*.

The second catalytic function concerns identity. LLMs do not normally articulate ideas from a stable, pre-existing personal identity. They do not possess a continuous personal history in the same way a human speaker does. Instead, they generate language that projects a speaker on the basis of linguistic cues in the prompt and patterns in their training data.

This ability shows that identity can be partially performed through language. At the same time, it reveals the social structures embedded in language itself. Training data contain social norms, exclusions, patterns of representation, and structural biases. As Bender and her coauthors argue, language models can reproduce and amplify patterns contained in their training data.[10] Feenberg likewise reminds us that technical systems are not neutral: their design and use embody social choices and values.[11] The LLM therefore reflects not only what has been said but also which forms of language have been most strongly represented.

In this limited sense, the LLM becomes a mirror of cultural language. It can reveal recurring patterns in how identities, authorities, and social categories are represented. This does not mean that it gives us direct access to a collective unconscious. Rather, it makes linguistic patterns visible and therefore provides material for further human interpretation.

5. Cognitive Risk and Human Re-encoding

The hallucinations of LLMs—their fluent but false or malformed utterances—show

---

[10] Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery, 2021), 610–23.

[11] Andrew Feenberg, *Questioning Technology* (Routledge, 1999).

how easily linguistic fluency can separate from truth and stable semantic attachment. Yet the wider concern is not simply that LLMs sometimes produce false statements. Their growing success may also change the habits through which humans produce and interpret language.

As language generation becomes cheaper and faster, linguistic output can increasingly function as a behavioral interface: its value may be judged by how efficiently it produces an answer or reaction rather than by how carefully its meaning has been developed. We may describe this tendency as movement toward a more *glyphic culture*, in which linguistic signs are increasingly consumed as rapid units of information and response. This does not imply the disappearance of meaning, but it may reduce the amount of effort devoted to its stabilization.

The metaphor of *Optimal Brain Damage* provides one way of expressing this concern. In machine learning, LeCun, Denker, and Solla use the term for the strategic pruning of connections in a neural network to improve efficiency.[12] Applied to human cognition, the phrase should remain an analogy rather than a literal neurological claim. As people increasingly outsource memory, analysis, and even sentence production to LLMs, they may have fewer occasions to practice some of the demanding cognitive tasks that writing traditionally requires.

Recent research gives some reason to investigate this possibility, although it does not prove cognitive decline. Kosmyna and her coauthors report differences in neural connectivity, recall, and reported ownership among participants writing essays with LLM assistance compared with other conditions.[13] These findings do not show that AI literally "rewires" the brain in a damaging way, but they raise a relevant question about how different tools alter the cognitive process of writing.

John Warner's argument in *More Than Words* helps clarify what may be at stake. Warner treats writing not simply as a method for recording already completed thoughts but as a process through which thought develops.[14] Writing requires the writer to discover relations, confront uncertainty, and decide what a sentence actually commits them to. If AI allows users to generate finished language before undergoing this process, the danger is not merely that

---

[12] Yann LeCun, John S. Denker, and Sara A. Solla, "Optimal Brain Damage," in *Advances in Neural Information Processing Systems 2* (1989), 598–605.

[13] Nataliya Kosmyna et al., "Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task," preprint, arXiv:2506.08872, 2025.

[14] John Warner, *More Than Words: How to Think About Writing in the Age of AI* (Basic Books, 2025).

the text is less authentic. The more important risk is that linguistic production becomes separated from conceptual formation.

This danger appears in both writing and reading. During writing, AI can produce language without requiring the user to make every conceptual choice. During analysis, AI can reduce a complex text to a set of points and therefore remove some of the difficulty through which interpretation develops. Efficiency is not itself the problem. The question is whether efficiency replaces the intellectual resistance that forces a user to stabilize meaning.

6. Adversarial Inquiry

Confronted with this possibility, we must ask what intellectual practice can prevent AI-assisted language from becoming passive consumption. I propose an *Adversarial Inquiry Approach*, in which the user takes the role of an examiner and debater.

Adversarial Inquiry means systematically interrogating LLM outputs by questioning their premises, demanding evidence, providing counterarguments, and testing whether the same concepts are being used consistently. When the model transfers a term from one field to another, the user asks whether the original implications remain valid. When a claim sounds persuasive, the user asks which inferential steps actually support it.

This process should not be described as literally reproducing reinforcement learning. An individual user is not technically retraining the model. The important effect occurs in the user's own reasoning. By challenging the model's output, the user is forced to identify and articulate the rules that stabilize the relevant concepts.

Adversarial Inquiry therefore works as a method of human re-encoding. It turns the user from a passive consumer of fluent text into an active participant who reconnects linguistic output with evidence, context, and conceptual relations. In terms of SAL, the user asks what would be required to move an output from weak semantic attachment toward a more strongly embedded use.

This is also why the model's imperfections can become intellectually productive. A weak or misplaced connection may force the human user to explain why two concepts should remain distinct. The model does not simply give the user knowledge; it provides propositions against which the user can test their own judgment.

7. Conclusion

This paper has argued that LLMs can generate stable linguistic form while meaning

becomes less stable when their utterances move across contexts without sufficient attachment to practice. PSC and SAL provide a framework for analyzing this process. PSC identifies the portable semantic core that allows a word to remain recognizable. SAL describes the degree to which a particular use is embedded in practical rules and conceptual networks. Semantic drift occurs when the first survives but the second weakens.

This framework suggests that LLMs should not simply be described as systems without meaning. They inherit substantial semantic structure from the human language on which they are trained. Yet they also do not independently determine all of the context-specific conditions that stabilize meaning. Their characteristic role is therefore catalytic: they generate linguistic and conceptual possibilities that require human interpretation and re-encoding.

Adversarial Inquiry offers one response to this condition. By treating model output as a proposition to be examined rather than a conclusion to be accepted, users can preserve the process of judgment that gives semantic relations their stability. The goal is not to eliminate AI from intellectual activity but to prevent fluency from being mistaken for completed thought.

This issue also connects with broader theories of cognition. Clark and Chalmers' idea of the *extended mind* suggests that cognitive outsourcing does not necessarily amount to cognitive decline. External tools can become parts of larger cognitive systems.[15] LLMs may therefore extend human capacities rather than simply replace them. Hayles's account of *cognitive assemblages* develops an even more distributed picture in which human and nonhuman actors participate together in cognitive processes.[16] If writing increasingly becomes a co-created process, authorship and agency may also become more distributed.

Walter Ong's account of *secondary orality* provides another way of thinking about this transformation. Ong describes electronic media as recovering some of the immediacy and participation of oral culture while remaining dependent on writing and print.[17] AI-mediated language may extend this development by making written language increasingly interactive and conversational. We might cautiously describe this as a form of "tertiary orality," although

---

[15] Andy Clark and David J. Chalmers, "The Extended Mind," *Analysis* 58, no. 1 (1998): 7–19.

[16] N. Katherine Hayles, *Unthought: The Power of the Cognitive Nonconscious* (University of Chicago Press, 2017), 203.

[17] Walter J. Ong, *Orality and Literacy: The Technologizing of the Word*, 30th anniversary ed. (Routledge, 2012), 132.

the term should remain a hypothesis rather than a settled historical category.

The central question, however, remains the same. When language can be generated with unprecedented speed, who performs the work that stabilizes what those words mean? If human users continue to question premises, test concepts against practice, and accept responsibility for the meanings they adopt, LLMs can become powerful tools. If linguistic production replaces these practices, then the deeper danger is not simply hallucination but the separation of language from the intellectual activity through which meaning becomes stable.

The AI revolution therefore raises a problem not only about what machines can do but also about what forms of thinking humans choose to preserve. When writing becomes increasingly collaborative between human and machine, the important question is no longer only who produced the words, but who remains responsible for making those words meaningful.

Bibliography

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, 2021.

Bender, Emily M., and Alexander Koller. "Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2020.

Clark, Andy, and David J. Chalmers. "The Extended Mind." *Analysis* 58, no. 1 (1998): 7–19.

Feenberg, Andrew. *Questioning Technology*. Routledge, 1999.

Hayles, N. Katherine. *Unthought: The Power of the Cognitive Nonconscious*. University of Chicago Press, 2017.

Hofstadter, Douglas R. *I Am a Strange Loop*. Basic Books, 2007.

Kaplan, Jared, Sam McCandlish, Tom Henighan, et al. "Scaling Laws for Neural Language Models." Preprint, arXiv:2001.08361, 2020.

Kosmyna, Nataliya, Eugene Hauptmann, Ye Tong Yuan, et al. "Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task." Preprint, arXiv:2506.08872, 2025.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962.

Lakoff, George. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press, 1987.

LeCun, Yann, John S. Denker, and Sara A. Solla. "Optimal Brain Damage." In *Advances in Neural Information Processing Systems 2*, 598–605. 1989.

Ong, Walter J. *Orality and Literacy: The Technologizing of the Word*. 30th anniversary ed. Routledge, 2012.

Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–57.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, et al. "Attention Is All You Need." In *Advances in Neural Information Processing Systems 30*. 2017.

Warner, John. *More Than Words: How to Think About Writing in the Age of AI*. Basic Books, 2025.

Wittgenstein, Ludwig. *Philosophical Investigations*. Rev. 4th ed. Translated by G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte. Wiley-Blackwell, 2009.

Research Paper

The Attribution Default After LLMs

Letian Cheng
Grade 12

Editors' Introduction

This paper asks on what evidence we may responsibly attribute understanding to large language models. It names a long-standing human habit the External Attribution Default: treating outward behavior as normally sufficient grounds for attributing mental states. The paper argues that this default reflects the limited evidence humans usually have about one another, not a general truth about attribution. With LLMs, reasoning summaries, tool traces and interpretability methods make evidence about *how* a response was produced increasingly available. Drawing on a deliberately weak Total Evidence Requirement and two well-chosen cases, the paper concludes that external behavior no longer counts as sufficient by default for LLMs. The cases are a model that copies its answers from an answer key and a parole-decision model that tracks race internally.

The paper's notion of "internal evidence" covers both what a component causally contributes to a response and what it represents, and the paper does not separate the two. It also chooses a weak version of the Total Evidence Requirement without comparing it to stronger epistemic principles, so the reasons that weaker version is enough are left largely implicit.

1. Introduction

An increasingly discussed question in philosophy of mind and AI is whether understanding can be attributed to AI systems.[1] Many current debates on this topic focus mainly on the model's outward performances and external evidence. Despite their differences, many of them give a central role to outward behavior or publicly available evidence. Cappelen and Dever's Observation Premise in their book *Going Whole Hog*[2] shares a similar spirit. The premise goes like this: "LLMs, e.g., ChatGPT, can understand the questions we ask it, and it can answer our questions. ChatGPT can also make suggestions, correct its mistakes, tell jokes, and provide arguments for what it says."[3] They defend this premise by appealing to an analogy between human and LLM opacity "Humans, like LLMs, are opaque black boxes, and in some sense it's incredibly hard to make sense of what's going on inside them. . . . Why should we expect things to be different for LLMs?"[4]

I agree that the Observation Premise provides us with a good starting point. However, their argument also exposes a human habit of inferring attribution from external evidence. In everyday life, we don't need to look inside each other's heads to know that they understand, and sometimes it is hard to do so. I call this habit the External Attribution Default:

External Attribution Default

> for humans, the historically stabilized and contingent practice of treating ordinary external observations as the default and normally sufficient basis for attributing mental states to others.

By "sufficient," I mean sufficient as an epistemically responsible stopping point, especially when the attribution is used in philosophical or high-stakes contexts.

I also distinguish internal evidence and external evidence. Evidence must be accessible to function as evidence, but the item the evidence is about can be external or

---

[1] Cf. A. M. Turing, "Computing Machinery and Intelligence," *Mind* 59, no. 236 (1950): 433–60, https://doi.org/10.1093/mind/LIX.236.433; J. R. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24, https://doi.org/10.1017/S0140525X00005756; and M. Mitchell and D. C. Krakauer, "The Debate over Understanding in AI's Large Language Models," *Proceedings of the National Academy of Sciences* 120, no. 13 (2023): e2215907120, https://doi.org/10.1073/pnas.2215907120.

[2] H. Cappelen and J. Dever, *Going Whole Hog: A Philosophical Defense of AI Cognition*, preprint, arXiv:2504.13988v1, 2025, https://doi.org/10.48550/arXiv.2504.13988.

[3] Cappelen and Dever, *Going Whole Hog*, 20.

[4] Cappelen and Dever, *Going Whole Hog*, 27.

internal. By external evidence, I mean evidence about the system's final outward performance: its answer, action, joke, correction, or argument. By internal evidence, I mean evidence about the production of that performance: the states, intermediate steps, traces, tools, or representations involved in producing the final response. Some of these items are mechanistically internal, such as activations and hidden features, which are directly inside the model's architecture. Some may be intermediate, such as reasoning tokens, tool-use results used before producing the final answer, which can be more easily understood. Some are indirect, such as summary of the original raw CoT. The above evidence differs, but they can all be considered internal evidence since what they are about is the model's internal production, not the external response. The distinction concerns what the evidence is about, not where it is shown.

To have a more concrete understanding, consider a human analogy. For humans, internal can be what occurs "under your skin" before you state your response, e.g. your brain's reasoning process, neurons firing. For LLMs, it should be what occurs "under the screen" of the machine. A human may use a detector on his head to detect neural signals, but the neurons' firing remains an internal process.

My concern is about how attribution should proceed once LLMs make new kinds of evidence available. In the human case, this default has been a fact throughout our history: humans never had any simple observations of each other's neurons' firing, electric signals flowing in the brain. However, this habit was conditioned by the ways of observation available to humans in history. For modern and future LLMs, the situation differs significantly. Providers can increasingly determine how much internal evidence can be available to users. Consider chain-of-thought as an example: The reasoning trace summary of an LLM can be hidden or shown, and such changes are simple and accessible. The central question is: once relevant internal evidence becomes ordinarily accessible, can the External Attribution Default still retain its status as a sufficient stopping point?

The paper proceeds as follows. I begin by identifying a habit that has been stable in human history. In section 2, I explain why the External Attribution Default has been so stable and natural for humans, and analyze some theories. Section 3 establishes the Total Evidence Requirement and presents the key argument. Section 4 shows that some internal evidence is now ordinarily available, using chain-of-thought as the basic example. Section 5 establishes the relevance premise by showing that internal evidence can change our confidence and belief upon an attribution made from external observations. I also reply to two objections in this section: Cappelen's skepticism about mechanistic interpretability, and the objection that

my argument is only a trivial total-evidence point. Section 6 draws two epistemic consequences and concludes.

2. The Human Evidential Condition

The External Attribution Default is best understood as a natural and reasonable method in the human world. We apply this default every day in the human world, but often without any notice as it is so natural and reliable for humans. There are two potential reasons.

The first reason is biological. Our biological structures throughout history rarely allowed ordinary internal evidence. Before a human speaks or responds, the relevant internal processes are mostly neural or psychological processes that take place in the brain. Throughout ordinary human history, it has been incredibly hard to look inside our brains. Hence, one cannot ordinarily look under each other's skin in conversations and examine their thoughts at the moment. Humans can still report their inner thoughts to others, but our habit of having conversations rarely requires us to do so, and it would be quite inconvenient. External observations appeared to be the most natural and accessible evidence. Internal evidence was simply not ordinarily available to humans. The Total Evidence Requirement still applies, but ordinarily accessible human evidence contained little or no internal evidence. Internal states were not considered unimportant. They were simply inaccessible. The choice of external evidence as default was implicitly forced by the biologically stabilized world for humans, not made.

The second reason is practical. As Cappelen and Dever state in their book, humans are indeed good detectors of mental states.[5] We can tell if someone understands a question by looking at her answer, facial expression, explanation, response to further questions. As humans, we know how ordinary humans tend to act when they have a certain mental state. For humans, external evidence is reliable enough for everyday attribution. Ordinary human social life has trained us throughout our history, and it wouldn't be a surprise that we learned to use such evidence with considerable success. Another practical point is that mental attribution is often a public practice. Its evidence must be shareable and comparable among speakers to be used. External behavior naturally satisfied this role in the human case.

Many theories developed around human mentality have given a central role to behavior, public criteria, or publicly observable evidence, which reflects how the External Attribution Default has influenced humans. I use these approaches only to illustrate the

[5] Cappelen and Dever, *Going Whole Hog*, 28.

naturalness of the default, and I do not object to them. Ryle provides a strong version of this orientation. In *The Concept of Mind*, Ryle argued that mental terms are descriptions of how an agent "did, or would, behave in various situations," not reports of "hidden inner episodes."[6] For humans, abilities and dispositions are mostly or nearly always judged through external behaviors and actions. Wittgenstein's approach is radically different; he was not a behaviorist in any explicit sense. However, his remark in the *Philosophical Investigations* that "an 'inner process' stands in need of outward criteria"[7] captures a similar background in a different sense. For mental-state attributions to have a place in public language, they must be connected with criteria that can be shared, applied, evaluated in public practice. For humans, these criteria are usually outward criteria: speech, action, response, expression, and participation in public forms of life. Davidson starts from another basis. In "Radical Interpretation," the interpreter begins from the "observed behavior of speakers in conjunction with observation of the basic perceptual stimulations that give rise to that behavior," and proceeds under the demand to "optimize agreement between ourselves and those we interpret."[8] In this view, attribution is a practice carried out in a public space that maximizes coherence. Again, the point is that, for human interpreters, the ordinary way of attributing states for other minds has been through external behavior in a shared world. Cappelen and Dever's Observation Premise captures a similar spirit. It shows how natural it is to transfer the human External Attribution Default to artificial systems: if external observation is enough for human attribution, why should LLMs be different? They also concede that simple observations may start the discussion without ending it.

These examples help explain why the External Attribution Default is so natural. For human mental attribution, internal evidence was rarely available, while public behavior was accessible, shareable, and reliable enough for everyday purposes. The success of the default depended on the human evidential condition, not that external evidence is always sufficient. This habit is useful and reliable for humans. However, my point is that its ordinary success depends on evidential conditions that change in the LLM case.

---

[6] G. Ryle, *The Concept of Mind* (Hutchinson, 1949).

[7] L. Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe (Blackwell, 1953), §580.

[8] D. Davidson, "Radical Interpretation," *Dialectica* 27, no. 3–4 (1973): 313–28, https://doi.org/10.1111/j.1746-8361.1973.tb00623.x.

3. The Total Evidence Requirement

To make an attribution, we need evidence. For humans, we mostly need external evidence as it is reliable enough. But more generally, we need to clarify how available evidence should be used.

I begin by providing a weak requirement:

The Total Evidence Requirement: To make an epistemically responsible conclusion, one should not ignore relevant and available evidence.

By "relevant evidence," I mean evidence that gives us a reason to change our conclusion or our confidence in it. It may either support or undermine the conclusion. By "available evidence," I mean evidence that is already ordinarily accessible and can be reasonably analyzed by humans, which may vary depending on different contexts.

Normally, if the evidence is both relevant and available, ignoring it would undermine your conclusion's ability of tracking the truth. To be epistemically responsible, we want our conclusion to track the truth. Also, denying this principle would seem to make your conclusion arbitrary, as you may simply consider only the evidence that you favor. I treat this as a minimal epistemic requirement. Its force comes from being applied to a new evidential situation. In the human case, it rarely required us to consider internal evidence, because such evidence was not ordinarily available. In the LLM case, this condition changes.

The key argument of this paper follows:

P1. *Observation:* Some evidence about LLM internal production(internal evidence) is ordinarily accessible.

P2. *Total Evidence Requirement*: To make an epistemically responsible conclusion, one should not ignore relevant and available evidence.

P3. *Relevance*: Some internal evidence is relevant to attribution.

C1. When relevant internal evidence is available, external evidence alone can no longer function as an epistemically responsible stopping point.

Section 4 establishes P1 and section 5 argues for P3. Consequently, the status of the External Attribution Default would be different in the LLM case. For humans, the default could function as normally sufficient because we could not ordinarily use internal evidence and external evidence has been reliable enough. For LLMs, the condition changes as stated by P1. From C1 it follows that the External Attribution Default loses its default status when relevant internal evidence is available.

4. Internal Evidence in LLMs

This section establishes P1: Some LLM internal evidence is ordinarily accessible. As I said before, internal evidence here requires only that the evidence is about the internal production.

Why does P1 hold for LLMs but not current humans? This is because, in principle, LLMs are very different from humans in terms of observation. LLMs have various interesting characteristics: they can be formulated as pure mathematical models. Engineers can observe the exact state of weight matrices by freezing the model at inference time and then inspect the tensor files; though they may be of vast size and we may not understand them clearly now. Recent techniques such as Natural Language Autoencoders[9] also illustrate the same trend: some model-internal activations can be translated into natural language explanations, although such explanations remain imperfect. When the model is open-sourced, we may gather sufficient information to create another identical LLM. In contrast, for humans the above seems relatively hard. Although the process may be possible in the future through BCI technologies, the current observations available to us remain relatively limited.

Before going into the concrete example, I will first brief explain what Chain-of-Thought is. A Chain-of-Thought (CoT) trace is a sequence of intermediate tokens generated before the final answer. These tokens are fed back into the context and affect later generation. Chain-of-Thought may go like this: To analyze the problem, I need to focus on the left corner of the picture . . . Chain-of-Thought usually represents the step-by-step intermediate steps produced.[10] A human analogy of Chain-of-Thought: when asked to make a choice in exams, your inner thoughts or thinking process regarding the question's conditions, methods of calculation. What you finally write down is "The answer is A."

In practice, more and more evidence about internal production is already available. Chain-of-Thought is the most accessible and ordinary illustration of this. Nowadays, users can ordinarily see the shortened version, a provider-generated summary or even the original trace. Raw CoT tokens causally participate in the internal production of the response. Though summaries and traces shown to the users differs from raw CoT tokens, they are still evidence

---

[9] K. Fraser-Taliente et al., "Natural Language Autoencoders Produce Unsupervised Explanations of LLM Activations," *Transformer Circuits Thread*, May 7, 2026, https://transformer-circuits.pub/2026/nla/.

[10] J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," in *Advances in Neural Information Processing Systems 35* (2022); T. Kojima et al., "Large Language Models Are Zero-Shot Reasoners," in *Advances in Neural Information Processing Systems 35* (2022).

about the internal production, which makes these summaries and traces internal evidence. Many current LLM systems choose to expose reasoning summaries, thinking traces, tool-use records, or CoT-like materials. For example, OpenAI states that raw reasoning tokens are not exposed but reasoning summaries can be requested; Anthropic's Claude API can return "thinking" content blocks when extended thinking is enabled; Google's Gemini API supports thought summaries.[11] When exposed, CoT-like content can be read, quoted, and evaluated by users. This shows that some internal evidence can be ordinarily available to us through interface choices, and whether the user can see the raw CoT is not the key here.

Two objections arise here:

1. Chain-of-thought is simply another output of the model as it is also tokens generated by the model, which makes it external evidence. Humans can also report their thoughts while they reason.

2. Chain-of-thought may have little to do with the model's internal states; it may be only a post-hoc, inaccurate explanation.

For the first objection, I agree that CoT consists of generated tokens and is therefore an intermediate output, since it can be "printed out." However, this does not show that CoT is not internal evidence. CoT's role in generation differs from the final answer as it is fed back into the context for further generation, which produces the response we see. From a functional point of view, CoT directly contributes to the internal production process, which matches our definition. Whether users are allowed to see a summary of them can be a dynamic interface choice, but that doesn't change that they are still about the model's internal production.

For the second objection, I agree that Chain-of-thought text may not accurately represent the causal process that produced the output. There is also a growing body of empirical evidence suggesting that CoT faithfulness varies significantly across model types and tasks,[12] just like humans' reported reasoning may not be representing the exact neural

---

[11] OpenAI, "Reasoning Models," OpenAI API documentation, accessed July 1, 2026, https://developers.openai.com/api/docs/guides/reasoning; Anthropic, "Building with Extended Thinking," Claude API documentation, accessed July 1, 2026, https://docs.anthropic.com/en/docs/build-with-claude/extended-thinking; Google, "Gemini Thinking," Gemini API documentation, accessed July 1, 2026, https://ai.google.dev/gemini-api/docs/thinking.

[12] M. Lanham et al., "Measuring Faithfulness in Chain-of-Thought Reasoning," preprint, arXiv:2307.13702, 2023, https://doi.org/10.48550/arXiv.2307.13702; M. Turpin et al., "Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting," in *Advances in*

process. However, evidence can often be observations limited by the methods available to us, but they still contribute to the conclusion we draw. Many evidential sources are imperfect or noisy, but the inaccuracy of CoT does not prevent us from gathering useful information out of it. I do not claim CoT reveals the true internal reasoning mechanism of the LLM. What CoT-like content really shows is the availability of some internal evidence, and that is enough for our argument here.

The above only shows that internal evidence can be available, but why should we consider it? I will explain this with two cases in section 5.

5. Why Internal Evidence Is Relevant

In section 2, I defined evidence as relevant when it gives us a reason to change a conclusion or our confidence in it. In this section, I establish P3, the relevance premise. The first case concerns cognitive attribution directly. The second extends the same point to a normatively important attribution.

Consider a model being tested to solve hard mathematical problems. It has given all correct answers and detailed rigorous explanation that cannot be distinguished from a human expert. It applied the right theorems, used the right tools, and followed appropriate logic. Observing the model's external output, it seems natural to attribute understanding of the problems to the model. Now suppose that we have discovered some internal evidence. The evidence may be raw CoT tokens or tool-use logs. It shows that the model solved the problems in a different way. Instead of trying to reason to solve it, it identified the problems as a benchmark, searched in a local answer-key dictionary, found the correct option and an expert explanation written by a human, and then simply copied it. This undermines the support that the perfect external performance gives to the attribution of understanding. A correct answer produced by solving a problem is epistemically different from copying a human expert's answer. The point only requires that this difference can change our confidence in the attribution. If one insists that answer-key search and problem-solving make no difference at all, then one is using a thin notion of understanding very close to reliable input-output success. Since internal evidence changes our confidence in the attribution, it is relevant.

The second case concerns a different attribution: whether a model understood and followed a normatively important instruction. I will first explain some MI techniques. Layer-

Neural Information Processing Systems 36 (2023).

wise probing tests whether some information can be identified from the model's activations at different layers. Causal intervention provides stronger evidence. Researchers change, remove, or replace only a selected activation, and then examine whether the model's output also changes. Consider an LLM designed to automatically make parole decisions. This is a hypothetical case modeled on existing decision-making algorithms and probing/intervention methods..[13] The model was provided with detailed accounts of the applicant's crime, performance in prison, application history and other records. However, race-related elements are not listed as the input variables and the model is explicitly required not to consider any race-related elements in its decision making.

First, we may use different evidence to make an attribution. At the external level, everything appeared to be perfectly fine. The model uses detailed reasoning and follows the instructions exactly such that its output cannot be distinguished from an expert's assessment. From external evidence alone, it appeared that the model understands the requirement and has produced a reasonable decision. Now suppose that additional internal evidence was obtained through layer-wise probing and causal intervention.[14] One important fact is that although race is not given as an explicit variable, the model may form a race-tracking internal feature from relevant information such as names, locations, or other details in the record. We used a layer-wise probe to identify race-related information from the model's intermediate representations in numerous layers. The result showed that race-related features are present. We then used a causal intervention. By directly changing or suppressing only this race-related internal feature at inference time, the parole decision changed systematically. This gives us evidence that the race-related feature played a causal role in producing the output.

Now compare the attribution we made from different evidence. Based on external evidence only, we would have the conclusion that this decision and explanation manifest the model's understanding and application of the instruction. Once the internal evidence is added, at least three interpretations become possible. The model may have failed to understand the instruction. It may have understood the instruction but failed to follow it. Or it may have

---

[13] Cf. Z. Obermeyer et al., "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," *Science* 366, no. 6464 (2019): 447–53, https://doi.org/10.1126/science.aax2342; J. Angwin et al., "Machine Bias," *ProPublica*, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; and J. Dressel and H. Farid, "The Accuracy, Fairness, and Limits of Predicting Recidivism," *Science Advances* 4, no. 1 (2018): eaao5580, https://doi.org/10.1126/sciadv.aao5580.

[14] N. Elhage et al., "A Mathematical Framework for Transformer Circuits," *Transformer Circuits Thread*, 2021, https://transformer-circuits.pub/2021/framework/index.html.

understood the instruction while giving an unfaithful explanation. The internal evidence does not fully determine which interpretation is correct. However, under each interpretation, the original conclusion drawn from the perfect response that the model understood and applied the instruction must be reconsidered. Consequently, internal evidence is relevant in this case.

Cappelen and Dever raise worries about mechanistic interpretability in their book.[15] Their concern is that even if MI uncovers a causally active race-tracking internal feature, it does not show that the specific feature has the meaning (e.g. race) we assigned to it. A feature visualization, for example, may be labelled as a dog-snout detector by us but not necessarily show that the feature really represents the actual concept of dog-snout; Applied to the parole case, their objection is that a probe or intervention may reveal a race-tracking internal feature, but this cannot show that the feature represents race, or that the model used race in its reasoning.

I do not need MI to recover full representational content. My argument only requires the evidence to be relevant to an attribution. If we found an internal feature that systematically tracks race across controlled cases, and intervening on it changes the parole decision, then our confidence in the original attribution should change as this feature is both closely related to race and important for the decision. We may not know how this feature is connected with our concept of "race," but the evidence surely makes it harder to believe that the model understood and followed the instruction not to use race. The remaining possibilities include misunderstanding, non-compliance, and an unfaithful explanation. Still, it is less likely that the model understands and actually follows the instruction compared to the attribution made from only external observations, and that is enough for relevance.

A more general objection is that the argument seems weak. Even if we accept internal evidence can be a defeater sometimes, then this does not influence external observations as the default basis.

I concede that the requirement seems weak, but the argument's importance is actually illustrated when internal evidence can be a defeater. As humans, we have taken the External Attribution Default as a habit, and it's both hard and unnecessary to change that in most conditions. The Default may still function as a good starting point in ordinary life. However, in contexts that require us to be epistemically responsible, we must not take external evidence as automatically sufficient. External evidence may still be useful, but we must not take it as the only useful evidence. This clarifies the importance of the argument: human attribution

[15] Cappelen and Dever, *Going Whole Hog*, 103.

practice was biologically and practically determined by the evidence available to us, while LLMs change what evidence can be made available through design choices.

Section 4 has argued for P1: evidence about LLM internal production is increasingly available. Some of it is already ordinarily accessible. Section 5 establishes that internal evidence is indeed relevant (P3), and C1 follows: external evidence is no longer sufficient when internal evidence is available. With the argument now complete, two consequences following from P1 and C1 can be drawn.

6. Consequences and Conclusions

One immediate consequence (following from P1) is that the evidence used in attribution of LLMs needs to be specified. Since what counts as observable evidence is a product of interface choices, any attribution of mental states depends on what providers make accessible and what users choose to consider. Providers can decide whether reasoning summaries, tool traces, retrieval logs are hidden, summarized, or surfaced. This means that two attributions based on the same response can have different epistemic standing if "the evidence considered differs. The attribution process is now dependent on two new variables: what providers choose to surface, and what users choose to consider.

The second consequence (following from C1) focuses on high-stakes contexts. The ethical weight is stronger in high-stakes contexts because the cost of relying on an attribution potentially defeated by available evidence is high. Considering only external evidence in everyday life is natural and reasonable since the epistemic risk is acceptable. However, in high-stakes contexts—parole decisions, medical diagnosis, legal adjudication—the cost structure reverses: simply observing externally may lead to injustice and undiscovered bias, and this penalty cannot be offset by the time saved by ignoring internal evidence. As we do need to be epistemically responsible in these cases, internal evidence should be carefully considered if available.

The External Attribution Default was stable in the human world. In that world, internal states did not ordinarily enter attribution practice. I have argued that this condition no longer holds for LLMs. The Total Evidence Requirement establishes that relevant and available evidence should not be ignored. The answer-key and parole cases show that internal evidence can be relevant. Since some internal evidence is now ordinarily accessible for LLMs, external evidence is not sufficient by default when relevant internal evidence is available.

My conclusion is limited. I have not argued whether LLMs understand or do not

understand, nor that internal evidence settles attribution alone. The argument in this paper is limited to current and future LLMs, although one could extend the argument to humans naturally. However, human neural data remains largely in laboratory settings and is less ordinarily accessible. If brain-computer interface techniques have developed so that human internal evidence becomes ordinarily accessible, the argument can be extended as the Total Evidence Requirement also holds for humans. For now, my conclusion is narrower: for current and future LLMs, the old External Attribution Default can no longer be taken without justification. We humans have formed this contingency under a series of biological and practical conditions, but we should be aware that attribution practices are shaped by evidential conditions. For LLMs, the evidential conditions change. The old external default can no longer function as an automatic stopping point.

Bibliography

Angwin, J., J. Larson, S. Mattu, and L. Kirchner. "Machine Bias." *ProPublica*, May 23, 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Anthropic. "Building with Extended Thinking." Claude API documentation. Accessed July 1, 2026. https://docs.anthropic.com/en/docs/build-with-claude/extended-thinking.

Bereska, L., and G. Bhatt. "Mechanistic Interpretability for AI Safety: A Review." Preprint, arXiv:2404.14082, 2024. https://doi.org/10.48550/arXiv.2404.14082.

Cappelen, H., and J. Dever. *Going Whole Hog: A Philosophical Defense of AI Cognition*. Preprint, arXiv:2504.13988v1, 2025. https://doi.org/10.48550/arXiv.2504.13988.

Davidson, D. "Radical Interpretation." *Dialectica* 27, no. 3–4 (1973): 314–28. https://doi.org/10.1111/j.1746-8361.1973.tb00623.x.

Dressel, J., and H. Farid. "The Accuracy, Fairness, and Limits of Predicting Recidivism." *Science Advances* 4, no. 1 (2018): eaao5580. https://doi.org/10.1126/sciadv.aao5580.

Elhage, N., N. Nanda, C. Olsson, et al. "A Mathematical Framework for Transformer Circuits." *Transformer Circuits Thread*, 2021. https://transformer-circuits.pub/2021/framework/index.html.

Fraser-Taliente, K., S. Kantamneni, E. Ong, et al. "Natural Language Autoencoders Produce Unsupervised Explanations of LLM Activations." *Transformer Circuits Thread*, May 7, 2026. https://transformer-circuits.pub/2026/nla/.

Goldman, A. I. "What Is Justified Belief?" In *Justification and Knowledge*, edited by G. Pappas. Springer, 1979. https://doi.org/10.1007/978-94-009-9493-5_1.

Google. "Gemini Thinking." Gemini API documentation. Accessed July 1, 2026. https://ai.google.dev/gemini-api/docs/thinking.

Kojima, T., S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa. "Large Language Models Are Zero-Shot Reasoners." In *Advances in Neural Information Processing Systems 35*. 2022.

Lanham, M., A. Chen, A. Rawal, et al. "Measuring Faithfulness in Chain-of-Thought Reasoning." Preprint, arXiv:2307.13702, 2023. https://doi.org/10.48550/arXiv.2307.13702.

Mahowald, K., A. A. Ivanova, I. A. Blank, N. Kanwisher, J. B. Tenenbaum, and E. Fedorenko. "Dissociating Language and Thought in Large Language Models: A

Cognitive Perspective." *Trends in Cognitive Sciences* 28, no. 6 (2024): 517–40. https://doi.org/10.1016/j.tics.2024.03.005.

Meng, K., D. Bau, A. Andonian, and Y. Belinkov. "Locating and Editing Factual Associations in GPT." In *Advances in Neural Information Processing Systems 35*. 2022.

Mitchell, M., and D. C. Krakauer. "The Debate over Understanding in AI's Large Language Models." *Proceedings of the National Academy of Sciences* 120, no. 13 (2023): e2215907120. https://doi.org/10.1073/pnas.2215907120.

Obermeyer, Z., B. Powers, C. Vogeli, and S. Mullainathan. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366, no. 6464 (2019): 447–53. https://doi.org/10.1126/science.aax2342.

OpenAI. "Reasoning Models." OpenAI API documentation. Accessed July 1, 2026. https://developers.openai.com/api/docs/guides/reasoning.

Quine, W. V. O. *Word and Object*. MIT Press, 1960.

Ryle, G. *The Concept of Mind*. Hutchinson, 1949.

Searle, J. R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24. https://doi.org/10.1017/S0140525X00005756.

Turing, A. M. "Computing Machinery and Intelligence." *Mind* 59, no. 236 (1950): 433–60. https://doi.org/10.1093/mind/LIX.236.433.

Turpin, M., J. Michael, E. Perez, and S. R. Bowman. "Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting." In *Advances in Neural Information Processing Systems 36*. 2023.

Wei, J., X. Wang, D. Schuurmans, et al. "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models." In *Advances in Neural Information Processing Systems 35*. 2022.

Wittgenstein, L. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Blackwell, 1953.

Research Paper

How Conversations Change Us: Supplying a Hermeneutic Mechanism for Dover's Conversational Self

Weizhong Li
Grade 12

Editors' Introduction

This paper takes up Daniela Dover's account of interpersonal inquiry, in which taking one another seriously requires the abdication of interpretive sovereignty: giving up the privilege of having the last word on oneself. Dover deliberately sets aside the question of how conversation changes one's self-understanding. The paper identifies this as a genuine gap in the literature and offers a novel answer. Drawing on Heidegger and Gadamer, it supplies a hermeneutic mechanism through which an interlocutor's words reshape one's self-understanding, and it recasts abdication as the foregrounding of one's prejudices and the turning of one's fore-structure toward oneself. Its application of Heideggerian "fore-sight" to the case Dover discusses is especially illuminating, and its positive account of how abdication works is sharp and persuasive.

The paper's critique of Dover rests on a particular reading of her "vetting process," one that attributes to her account commitments her text does not explicitly make. Its main contribution lies less in this critique than in its constructive hermeneutic account, which stands on its own as an explanation of how conversation reshapes self-understanding.

Editors' note: For clarity, the citations in this article have been converted to MLA format.

1. Introduction

Do people have privileged access to knowledge of themselves? Building on his famous statement "Ego cogito, ergo sum," Descartes argues that a thinking subject can have privileged and infallible access to her thoughts through introspection, while others' access is merely inferential (pt. 1, art. 7). Other philosophers relocate this privilege: on Gareth Evans's transparency account, privileged access to self-knowledge is grounded in looking outward at the world rather than inspecting an inner state (225). In contrast, a deflationary view of self-knowledge denies such privilege outright, since any self-ascriptions are themselves interpretations, as revisable and fallible as any interpretations from others (Carruthers 1-2). Daniela Dover, however, steps out of this metaphysical and epistemological debate about self-knowledge. Rather than arguing about whether introspection in fact grants privilege, she takes up a normative question: in our daily practices of interpersonal relationships, should one insist on such privilege?

Dover examines this question through what she calls interpersonal inquiry, conversations that are "animated by mutual curiosity" (193), which she argues is governed by an ethic of "taking one another seriously" (200). Normally, we, the interlocutors, insist on the presumed privilege of "having the last word when it comes to the subject of oneself," based on our perceived direct access to our own mental states. Dover calls this privilege *interpretive sovereignty*. In the ideal circumstance, we give up this privilege, fulfilling a condition Dover names *abdication of interpretive sovereignty* (202).

Dover is explicit that her focus on interpersonal inquiry is ethical. She is committed to an "ethics-first approach" in which she derives an ethics for conceiving selfhood and understanding from a phenomenological investigation into how we aspire to relate. She rejects deducing ethical concepts from metaphysical or epistemological theories and calls such methods a reversal of "both our explanatory and our ethical priorities," arguing that we "have a much firmer grasp" on the phenomenology of these experiences than on those theories (Dover 196). In her paper, she characterizes a way of conversing and presents the phenomenology of it in an attempt to show that it is worth wanting. Although she mentions that changes in one's self-understanding are likely results of this activity, she sets aside the specific hermeneutic process of how these changes happen. This essay's aim is to explain this process. To achieve this purpose, I will borrow Hans-Georg Gadamer's hermeneutic theories and show how it can supplement and enhance our understanding of Dover's account.

One might object that Gadamerian hermeneutics emerges downstream of an ontological conception of understanding, which Dover's ethics-first approach avoids. However, Gadamer similarly derives his hermeneutics from an examination of shared human interpretive practices (270). Additionally, investigating the hermeneutic aspect of conversation—how the interlocutors' utterances are interpreted and how they influence our understandings—can tell us whether a proposed ethical ideal is compatible with how interpretation actually works and therefore achievable.

Following Dover, I focus on self-understanding instead of self-knowledge. My aim is to present how one's interpretation of oneself might be altered during conversation, so whether that interpretation is correct is set aside. Knowledge, as suggested by Francey Russell, implies mastery and closure—inquiry pursuing knowledge is consummated when knowledge is obtained. This, in turn, can obscure the *autotelic* nature of interpersonal inquiry. Dover describes one's self-understanding as one's "working theory" of oneself and "the accumulated result" of one's efforts to understand oneself. Dover prefers it because it captures one's simultaneously "epistemic, ethical and aesthetic" relation with oneself, with a thicker sense than mere self-knowledge (197).

In section 2, I will start by reconstructing Dover's account of interpersonal inquiry to situate the abdication condition in her theoretical framework. Section 3 will argue that Dover is committed to a dubious interpretive apparatus of "a conscious, reflective vetting process," leaving how abdication can be achieved unexplained. Section 4 will argue that abdication can be recovered through Gadamerian hermeneutics as the foregrounding of *prejudice* and the turning of *fore-sight* toward oneself. Section 5 will argue that based on Gadamerian hermeneutics, changes to one's self-understanding cannot be predicted or traced, which Dover asserts but does not explain.

2. Dover's Account of Interpersonal Inquiry

Dover uses *interpersonal inquiry* to refer to conversations "animated by mutual curiosity." Notably, the inquiry should not aim at satisfaction of curiosity—it is *autotelic*, "pursued and valued for its own sake," not as a means to an end (Dover 194).

She distinguishes between two approaches to this kind of conversation. The first is what she calls the *revelation picture*, where the interlocutors take turns to disclose information about themselves. It reaches its purpose when both interlocutors "have not deceived" each other and have "succeeded in accurately perceiving what [each other has]

revealed" (Talbert 194, qtd. in Dover 199).

Dover points out that the revelation picture is mistaken in two ways. First, its account of self-revelation is too narrow, limited to "heart-to-heart confessions" of what was once concealed. However, in actual practice, mundane conversations about "gardening or food or dogs" can be as revealing, in the sense that "when I talk about myself, I am also talking about the world, and when I talk about the world, I give my interlocutor a sense of who I am" (Dover 199). Second, the revelation picture fails to capture the intersubjective nature of interpersonal inquiry. Ideal conversations should not be interspersed monologues, in which "one is cast as autobiographical narrator, the other as audience." Rather, the interlocutors' attempts to understand each other are "constitutively bound up" and "comprehensively intermingled." They do not take turns performing an activity but engage in one activity together (Dover 200).

Dover's alternative to the revelation picture involves "taking one another seriously." It requires both interlocutors to satisfy three conditions: input-seeking, abdication, and reciprocity (Dover 201).

The input-seeking condition is fulfilled when I take your self-understanding as relevant to my understanding of you. Simply accepting your point of view through mere observation from the outside is not enough, since it cannot answer "the questions that I will have in response to them." To meet this condition, "I have to actively interrogate your point of view from my point of view" (Dover 201).

The abdication condition means that I should take your understanding of me as relevant to my self-understanding. This requires the abdication of one's interpretive sovereignty, which is to give up "the privilege of having the last word when it comes to the subject of oneself" (Dover 203). Abdication does not mean transferring the privilege of the last word to the interlocutor and granting her the last word instead; nor does it demand merely outward politeness while I privately hold your words against "a conscious, reflective vetting process" before letting influence happen (203). That would be to treat your understanding of me as "mere evidence, or as mere suggestions to 'take or leave' depending on how they fare with respect to my own antecedent commitments, values and self-conceptions" (203). In order to abdicate one's interpretive sovereignty in the full sense, one has to open oneself to the "unpredictable and perhaps even untraceable" changes provoked by the conversation (203). As Dover summarizes, the abdication condition consists not in "the actual occurrence or course of such changes, but in the attitude to conversation that makes them possible in the first place, namely, a willingness to loosen my grip on myself by loosening my grip on my

antecedent self-understanding" (204).

Lastly, the reciprocity condition is met if I want and encourage my interlocutor to adopt the above two stances toward me: to take my self-understanding into account when she tries to understand me, and to let my understanding of her inform her self-understanding (Dover 204).

Dover holds that the reciprocity and input-seeking conditions are often easily satisfied for conversations between genuinely curious interlocutors (205). In contrast, the abdication condition is the most demanding one: we are often reluctant to give up our authority over a subject—ourselves—in which we have invested so much, since it's tempting to "control the stories people tell about us" (Dover 206).

3. The Vetting Process and Its Difficulties

In this section, I focus on the abdication condition because it directly addresses how one should treat the interlocutor's interpretation of oneself. I present the presumed interpretive apparatus is presupposes and argue that this presumption is untenable.

Dover uses interpretive sovereignty to refer to one's privilege to have the last word when talking about self-related topics that are "broadly humanistic, ethically 'thick,'" such as one's "antecedent commitments, values and self-conceptions." I will call the interlocutor's utterances about these topics the *ingredients of self-understanding*, abbreviated as *ingredients* for convenience. When describing how interpretive sovereignty can be abdicated, Dover discusses the existence of "a conscious, reflective vetting process" that should be suspended for the condition to be fulfilled (202–3). Its work is to preclude any influence from the interlocutor by rejecting the *ingredients* she provides when they do not correspond with my current self-understanding and only accepting them when they do. When the conscious vetting is not suspended, the utterances are treated as "mere evidence, or as mere suggestions to 'take or leave'" with little influence on one's self-understanding. In contrast, after the vetting is suspended, *ingredients* can influence self-understanding directly.

The vetting is described as a structure that the interlocutor's utterance engages after its meaning is settled but before it can influence one's self-understanding. In other words, Dover's description implies that there are three steps for us to interpret an utterance during the conversation: (1) make sense of the utterance/work out its meaning, (2) check it against our existing understanding, i.e., upholding the vetting, and (3) revise one's self-understanding. She implies that it is only possible for one's self-understanding to be changed if the vetting

process is suspended. In this way, she also presupposes that it is possible for the vetting process to be suspended.

However, the existence of such a "vetting process" is questionable. Dover's description attributes two properties—deliberate and conservative—to this process; I will argue that neither holds for how we actually interpret an utterance about ourselves.

(1) Deliberate. Dover considers the vetting process as not "a fully passive surrender" but a display of the "trumping authority of conscious reflection and deliberation" (203n14). In calling on people to suspend the vetting, Dover suggests that this could be achieved through individual deliberation, or it makes no sense to convince them to withhold the vetting. To summarize, Dover's abdication condition presupposes that through efforts of deliberation, people can suspend the cognitive process that evaluates interlocutors' utterances against their current self-understandings so that all *ingredients* are treated as important.

It is questionable whether this is possible. For a sentence about a topic to make sense and be understood, one must relate it to something one already knows. Since *ingredients* are about oneself, it is impossible to understand them without relating them to one's current self-understanding. Imagine a conversation between Dick and his mother. Dick's mother says to Dick, "You are too lazy. You never helped me wash dishes after lunch." This is an *ingredient of self-understanding* for Dick, since it makes a claim about what kind of person he is. As Dover's account implies, if Dick is willing to abdicate his interpretive sovereignty, he can understand this claim without being aware of its correspondence with his existing self-understanding.

Settling this requires knowing what Dick must already be able to do to understand what his mother said. Evans's account of how subject-predicate ascription is understood gives us a place to start. According to Evans, understanding "a is F" requires two components: an idea of the subject and a concept of the property. That is, if one can think "a is F," one must be able to think "a is G" for any G one grasps, and "b is F" for any b one can think of (Evans 104–6). For Dick to understand "You are lazy," two processes are proceeding simultaneously: First, he needs to understand that the claim is addressing *himself*, which requires a capacity to raise other self-ascriptions, in other words, a capacity to deploy his existing self-understanding. Second, he needs to be able to attribute the property "lazy" to other subjects, which is to comprehend the concept of laziness as a behavioral disposition involving the avoidance of work and effort. Nevertheless, this only fixes a category of behaviors, and the extent of the disposition remains unsettled. "Lazy" is, in the language of Graff, a gradable adjective whose predicative use attributes not a context-invariant property,

but a property related to norms that are interest-relative (64). In the case of Dick, the norm is an expectation that his mother placed on him. It consists of what is typical of a household member like Dick. More importantly, it is composed of his own commitment to the practice through his ordinary conduct. In this way, to identify the exact property he was attributed by his mother, Dick needs to at least have a sense of his "laziness-related" behaviors. Therefore, Dick cannot be totally unaware of how lazy he was in order to understand the utterance of his mother. To clarify, this means that Dick should be aware of the degree to which his mother's claim corresponds with his self-concept, so it does not matter if he and his mother might have different understandings of the semantic meaning of the word "lazy."

One might object that this conclusion might only apply to utterances about gradable properties. If Dick's mother makes a claim about a categorical character of his, such as "you are a high school student," then understanding this utterance does not entail an awareness of its correspondence to Dick's self-understanding. However, this utterance might not be regarded as an *ingredient of self-understanding* because it is a part of "workaday facts" that "can be settled with reference to public records." It only becomes an *ingredient* when it can be interpreted as "ethically 'thick'" (Dover 202), while these interpretations often connect the categorical character with gradable ones. For example, "you are a high school student" becomes ethically significant when it is interpreted as talking about whether Dick is hardworking enough as high school students are expected to be, or whether Dick loves doing sports as high school students typically do. Categorical properties start to gain ethical thickness when they are interpreted as referring to gradable ones.

Recall that Dover's account implies that a sense of how an utterance corresponds to one's existing self-understanding arises after one comprehends it. This step-by-step conception of interpretation is an oversimplification because evaluation and understanding are actually interwoven during our cognitive processing of an external claim. A sense of how consistent the *ingredients* are with one's current self-understanding arises together with the meaning of the utterance. Therefore, it is difficult for the evaluation, i.e., the vetting process, to be suspended by deliberation if one attempts to understand an utterance.

Some may refute that deliberation can help people escape the influence of the vetting process in another way: not by suspending the process itself, but by redirecting our attention elsewhere so that the process would stop automatically. This is feasible, but it cannot achieve the effect that Dover claims the abdication of sovereignty achieves. Dover argues that abdication of sovereignty "[opens] the door" for *ingredients* to let further influence happen that "takes place under the radar of *conscious* introspection" (203). Since the process of

evaluation is a part of the process of understanding, to redirect attention away from assessing *ingredients* would simultaneously terminate the process of understanding them. As a result, the *ingredients* would be ignored, kept off "the radar of *conscious* introspection".

Nevertheless, that the method that Dover provides for abdication fails does not mean that abdication itself is unachievable. Although the vetting process itself cannot be suspended by deliberative effort, influence of *ingredients* can still reach us. This is going to be demonstrated in the following examination of the vetting process's second property, its conservative nature.

(2) Conservative. In Dover's description, the purpose of the vetting process is to "[hold] out against any influence from you." This is achieved through checking the *ingredients* regarding "how they fare with respect to my own antecedent commitments, values and self-conceptions" (Dover 203). It operates like a filter that allows only what is in accord with one's antecedent self-conception or perspective to pass through. The function of vetting is to prevent one's current self-understanding from being changed by blocking anything that might conflict with or challenge it.

In the examination of the first property, I have demonstrated that the function of such a vetting process cannot be isolated from understanding. To show that this second property fails, I will assume that vetting is an independent process as Dover suggests. I will demonstrate that even though the vetting process is not suspended, self-understanding can change regardless. I argue for this claim in three cases: (1) the *ingredient* is in harmony with one's existing self-understanding, (2) the *ingredient* is not in harmony with one's existing self-understanding, and (3) the *ingredient* is about an aspect that one's existing self-understanding does not cover. I will continue to use the case in which Dick's mother told him he was lazy because he never helped her wash dishes. Assume that he is not willing to abdicate his interpretive sovereignty and suspend his conscious vetting.

Case 1. Dick is aware of the fact that he was lazy and admits that as a part of his self-understanding. Then his mother's utterance passes his vetting filter frictionlessly and comes to him as a restatement of a fact he already knows. To determine that his mother's claim is in harmony with his self-conception, he needs to consciously think about his existing self-conception to serve as a benchmark for comparison. According to Anderson and Schooler's model of memory, human minds do not make everything instantaneously available at the same time to optimize the use of cognitive resources. The availability of stored information is determined by losses of holding it, gains from successful retrieval, and the probability of a demand for retrieval (Anderson and Schooler 396–97). Applied to Dick, he does not think

about his self-conception all the time. Instead, it is stored in his memory most of the time and is only retrieved if cognitive tasks require it. This time, it is raised into consciousness because processing his mother's utterance needs it. This retrieval process can increase the availability of this self-conception by increasing the probability that the mind regards this information as being demanded. Therefore, although the content of Dick's belief that he is lazy hasn't changed, that self-conception becomes more salient.

Case 2. Dick does not think that he was lazy, so his mother's claim does not pass the vetting filter he upholds. Like *Case 1*, to reject his mother's claim, Dick first needs to establish that he is not lazy and decide that this claim does not align with his current self-understanding. In this process, he retrieves his existing self-conception, brings it to the level of conscious awareness, and in doing so reinforces it.

Moreover, since Dick's mother provides evidence for her claim that he was lazy—he never helped her wash dishes—he needs to offer alternative interpretations of the evidence she pointed out, or it contradicts the self-understanding he wants to preserve. Maybe he is not lazy but too tired to help her. Maybe he thinks it is wiser to be selective in choosing an activity to spend his time on. Maybe he helps her do laundry every day and thinks that his mother was too demanding in asking him to also wash dishes. Whatever his reinterpretation is, it becomes a modification to his current self-conception, which is also a type of change.

Case 3. Dick never thinks about whether he was lazy. Hence, to preserve his self-understanding, he may bring up relevant but not identical categories that already exist in his self-conception and check his mother's claim against them, such as whether he was responsible or whether he was productive. But no matter what benchmark he brings up, he eventually reaches some sort of conclusion on the topic of whether he was lazy. In this way, his mother's claim changes his self-understanding by making available a new category of self-evaluation to him.

To conclude, suspending the vetting process is therefore not a necessary condition for one's self-understanding to be changed by *ingredients* during conversations.

Recall that Dover locates the "conscious, reflective vetting process" as occurring after the meaning of *ingredients* is settled and before they can reach one's self-understanding to cause changes. In the above section, I demonstrated that evaluation of a claim's correspondence is intertwined with the process of understanding it. Then, I showed that alteration to one's self-understanding, similarly, is not separate but happens alongside the process of interpretation. As a result, there isn't such a structure as the vetting process between making sense of an utterance and changing one's self-understanding. One does

evaluate the correspondence of an *ingredient*, but that process is entangled with one's interpretation of it and its alteration to one's self-understanding. In short, Dover suggests that these three processes are separate, and that one of them can be eliminated, and I have argued against this possibility.

This does not mean, however, that abdication of sovereignty is unethical or unachievable. According to Dover, "[t]he abdication of interpretive sovereignty consists, not in the actual occurrence or course of such changes, but in the attitude to conversation that makes them possible in the first place" (203–4). The changes to one's self-understanding seem to occur regardless of one's attitude or whether interpretive sovereignty is abdicated. However, as Dover said, abdication is not defined by merely the existence of changes. It is defined by one giving up one's privilege to have the last word on oneself when encountering the *ingredients* provided by an interlocutor. At the same time, attitudes that Dover proposes as a part of abdication, such as "an openness to being changed" (203), are not meaningless or inconsequential. Although they are not necessary for changes to happen, they can still influence the range of changes that could occur and continue to be a vital part of Dover's ethical ideal of "taking one another seriously."

If abdication cannot be achieved by suspending the vetting process, then it must work by some other route to preserve Dover's ethical picture. Identifying that route requires an account of interpretation on which comprehension and evaluation are interwoven rather than sequential, and on which one's self-understanding changes regardless. The next section will show that Gadamer's hermeneutic circle supplies the interpretive mechanism behind Dover's ideal conversation.

4. Recovering Abdication in the Hermeneutic Circle

In this section, I will first argue that Gadamerian hermeneutics can be used to supply the mechanism of how one's self-understanding can be changed in Dover's "taking one another seriously." Then, I will recover abdication in this framework and explain how abdication can enable more thorough changes to one's self-understanding.

4.1 Gadamerian Hermeneutics Applied to Dover

Dover argues that abdication of interpretive sovereignty is ethical by illustrating its phenomenological appeal through the case of Hannah Upp, who exemplifies it. Upp is a woman with dissociative fugues, a pathological condition that makes her forget who she is

episodically, yet she "knew she was loving and openhearted." This makes giving up interpretive sovereignty especially easy for her and grants her extreme openness to changes in interpersonal interaction. She was an evangelical Christian when she met her friend Bhattacharya, who was raised Hindu. When Bhattacharya asked her, "Do you think I'm going to Hell?" she wept. Later, Upp was found to be a Quaker and dating a woman. She clearly underwent significant changes to her self-understanding. How did this change happen? Dover argues that Upp had abdicated her interpretive sovereignty since she "did not treat Bhattacharya's question merely as a first-order theological question to be answered by reporting on her antecedent beliefs, nor merely as a challenge to be met by defending them." Rather, Upp realized that answering "yes" to the question might imply "Upp herself as (for example) callous, arrogant or parochial" and recognized the question as relevant not only to her theology but also to her self-understanding (Dover 211). However, Dover's account of Upp's interpretive process is incomplete: she does not explain how a question about Bhattacharya's fate becomes a question about Upp herself, nor why Upp's religious belief, rather than her openheartedness, gives way.

This gap can be filled by an account of interpretation that explains the relationship between understanding the interlocutor's claim and changing one's self-understanding. Gadamerian hermeneutics, specifically Gadamer's concept of the *hermeneutic circle* and Heidegger's *fore-structure*, are suitable for this purpose for three reasons. First, despite explaining his theory through the interpretation of texts, Gadamer's hermeneutics is developed on a conception of interpretation that also takes conversation as a model of understanding (387). Second, Heidegger's distinction between *fore-sight* and *fore-having* is especially apt to show how the meaning of an utterance can change without the arrival of new information. Third, Gadamerian hermeneutics fulfills the requirement in section 2, that is, understanding the utterance, evaluating its fit with one's understanding, and changing one's self-understanding are intertwined together.

According to Gadamer's account of interpretation, understanding of a text begins with engaging with the text with "particular expectations" that lead to the emergence of the "initial meaning." Then, as this "initial meaning" is projected to the whole text, *fore-projection* is created and is then "constantly revised" according to the meaning that further emerges as one continues to read the text. The revision in turn reshapes the expectations that determine how meanings emerge. And this circular process of forming and refining *fore-projection* is understanding, which he calls the *hermeneutic circle* following Heidegger (Gadamer 269). Readers often approach the texts with an expectation that there's "an immanent unity of

meaning" (Gadamer 294), which partly explains why the purpose of understanding is to reach "the harmony of all the details with the whole" (291).

The "expectations" that Gadamer mentions are the concretization of Heidegger's *fore-structure* (298). Heidegger argues that interpretation is "never a presuppositionless apprehending of something presented to us," and the *fore-structure* is what understanding rests upon (191). It has three aspects: *fore-having* is the background "totality of involvements" and prior familiarity with the subject matter that one already brings in; *fore-sight* is the prior perspective that fixes the respect in which the thing is interpreted and makes one more attentive to certain aspects rather than others; *fore-conception* is the conceptual vocabulary already held in place for the interpretation (Heidegger 192). Gadamer argues that during interpretation, one is always going to expect that all the meanings are compatible with each other to form a coherent whole, which he called the *fore-conception* of completeness (294). Therefore, as further meaning emerges, *fore-structure* and *fore-projection* are revised continually until they reach inner coherence and cohere with each other.

In interpersonal inquiry, the object of understanding is the utterance one hears from the interlocutor. In this context, *fore-having* includes one's understanding of oneself, the interlocutor, and the world. *Fore-sight* is an anticipation of what the interlocutor's utterance is about, including its scope and illocution, which fixes how it relates to one's *fore-having*. These two aspects matter most for my purpose since fore-having holds self-understanding, while fore-sight determines whether an utterance is taken to be about oneself. Since self-understanding is part of *fore-having*, from which the projected meaning is formed, understanding an utterance about oneself already evaluates it against one's self-understanding; and when the projection fails to fit, self-understanding can be revised. Understanding, evaluation and change are therefore interwoven parts of one circle instead of three separate steps, fulfilling the requirements raised in section 3.

The same process is demonstrated in how Upp understood Bhattacharya's question "Do you think I'm going to Hell?" Based on an inference about her *fore-structure*, she considered herself to be "loving and openhearted" characteristic, recognized that Bhattacharya was her beloved friend, and had a worldview that's typical of evangelical Christians. Initially, her *fore-sight* anticipated that the question was about her first-order theological belief. She reached the new understanding that "I believe my friend should go to Hell," since, due to her religious commitment, those who aren't devoted to Jesus, including her Hindu friend, would go to Hell after death. Projecting this meaning to the whole conversation, she noticed the contradiction between her new self-understanding and her

recognition of Bhattacharya as her friend. A new *fore-sight* then formed—the question isn't about her religious belief but about her evaluation of friendship. So she began to understand Bhattacharya as asking, "Am I more important to you than your religious belief?" She wanted to answer yes, since she couldn't regard herself as "loving and openhearted" if she answered no. She then "turned the question in on herself" and treated it as related to her self-understanding (Dover 211). The question's meaning then becomes "Do I want to be a loving person at the cost of giving up my lifelong belief?" Her later changes indicated that she answered yes, recognizing "loving and openhearted" as more crucial to her self-understanding than her religious belief.

4.2 Abdication as Foregrounding of Prejudice

So far, I have used Gadamer to describe how self-understanding changes. The same framework also yields a more precise account of what the abdication requires. Dover describes the abdication of interpretive sovereignty positively as an attitude—"an openness to being changed" (203). In an ordinary reading of "openness," it requires one to empty oneself of prejudices one antecedently holds to accept the interlocutor's utterance in an unfiltered manner. However, in the hermeneutic circle, a comprehensive understanding of the interlocutor arises only through *fore-projection*, which is generated from one's prior involvement with the subject matter. Gadamer's name for the historically formed content of the *fore-structure* is *prejudice* (*Vorurteil*). Gadamer argues that it shouldn't be treated as the enemy of understanding, since "the fundamental prejudice of the Enlightenment is the prejudice against prejudice itself"; instead, prejudice is necessarily involved in all understanding as the very condition for it to happen (274). Gadamer defines prejudice as "a judgment that is rendered before all the elements that determine a situation have been finally examined" (273). Therefore, prejudice is not necessarily a false judgment but a judgment that's made at an early stage.

Then, if Dover's openness does not mean an absence of prejudice, what does it mean? According to Gadamer, openness "involves neither 'neutrality' with respect to content nor the extinction of one's self, but the foregrounding and appropriation of one's own . . . prejudices" (271). Foregrounding of a prejudice means that it is brought from operative-but-invisible to visible, while appropriation of it requires one to own one's own prejudice rather than dismiss its existence. Appropriation prevents foregrounding from being idle, since realizing one's assumption does not mean that one treats it as a prejudice instead of a plain fact. In other

words, openness requires one "to be aware of one's own bias, so that the text can present itself in all its otherness and thus assert its own truth against one's own [prejudices]" (Gadamer 272). For interpersonal inquiry, an operative but unnoticed prejudice provokes an interpretation of an utterance that is in harmony with itself, impeding any revision to itself. In contrast, once a prejudice is foregrounded and appropriated, it can stand in tension with the interlocutor's utterance: confirming, qualifying, contradicting, etc. As Gadamer claims, ideal interpretation is "not covering up this tension by attempting a naive assimilation" but "consciously bringing it out" (305). Therefore, even if "a conscious, reflective vetting process" exists, suspension of it would not help reach the purpose of abdication—enable more thorough changes to one's self-understandings—because it hides one's self-understanding and reduces its visibility. Instead, the open attitude stipulated by the abdication condition should put one's self-understanding at a stage where it becomes more obvious and thus susceptible to challenges.

4.3 How Abdication Makes a Difference

As the openness in abdication is redefined as the foregrounding of one's prejudices, it requires an explanation of why this openness is what "makes [changes to self-understanding] possible in the first place." Gadamer's hermeneutics can supply two routes through which abdication of interpretive sovereignty can make one's self-understanding more prone to change, depending on the kind of contradiction that drives the revision.

When an utterance contradicts one's self-understanding, the contradiction can be resolved in two directions: either to modify one's self-understanding or to reinterpret the utterance, i.e., either revise one's *fore-structure* or one's *fore-projection*. If one doesn't abdicate one's interpretive sovereignty and foreground one's prejudice, one's self-understanding is not put into question and remains invisible, so the contradiction is resolved largely by reinterpreting the utterance. Therefore, abdication makes self-understanding more available as what might be wrong and need revision.

When the utterance points out a contradiction between different existing self-understandings, as Bhattacharya's question did for Upp, the influence of abdication through the abovementioned route is not significant. At this point, reinterpreting the utterance cannot remove the contradiction within oneself, so self-understanding must be revised. Abdication influences this process by affecting whether the contradiction would be noticed.

If an utterance has the potential to expose a contradiction between multiple existing

self-understandings, abdication mainly exerts influence via the *fore-sight*. Among the three components of one's *fore-structure*, *fore-having* and *fore-conception* are things that I possess—neither is under deliberative control, and neither would be changed when one adopts a different attitude. In contrast, *fore-sight* fixes the respect in which things are examined and directs attention to certain aspects rather than others. It is a selection from different directions that the interpretation can lead toward. Therefore, even though the *fore-structure* is necessary for understanding and cannot be suspended, *fore-sight* opens possibilities for one's attention and attitudes to influence how things get interpreted.

One would inevitably change one's self-understanding during the conversation only when the interlocutor's utterance is understood as saying something about oneself. Therefore, one can prevent one's self-understanding from being changed by not relating an utterance to oneself. Although this avoidance is not possible when a claim has "you" as its grammatical subject, this mechanism can work when the interlocutor speaks about herself or other things. As a result, some utterances, despite being *ingredients of self-understanding*, might not get the chance to influence one's self-understanding because the person does not abdicate her interpretive sovereignty and directs her *fore-sight* away from herself.

For Upp, her friend Bhattacharya's question didn't address her directly—its direct subject is Bhattacharya's fate. Yet she "turned the question in on herself" and treated the answer to this question as related to what kind of person she is. With a willingness to be changed by her friend, her *fore-sight* gradually directed the focus of the question from other subjects to herself, eventually exposing a contradiction within her self-understanding. If Upp didn't abdicate her interpretive sovereignty, she would interpret Bhattacharya's question as revealing something else, maybe about mere theology, maybe about Bhattacharya's own insecurity, anything but her self-understanding.

In conclusion, abdication works in two ways: letting self-understanding, rather than the utterance's meaning, to bear revision when an utterance contradicts it, and affecting whether contradictions within one's self-understanding are exposed at all.

5. Unpredictability and Untraceability of Changes

Gadamer's framework can explain how the abdication condition can be achieved and why it has its intended effect. A further test of its adequacy is whether it can also explain the phenomena that Dover describes as likely to happen after interpretive sovereignty is abdicated. In this section, I will focus on how changes to one's self-understanding remain

unpredictable and untraceable.

Dover describes the ideal effect of abdication of interpretive sovereignty as "allowing myself to change in response to you in unpredictable and perhaps even untraceable ways" (203). She understood unpredictability and untraceability as separate traits. Unpredictability is prospective: no participant can foresee how their understandings would change, let alone control the direction of that change during conversations. In contrast, untraceability is retrospective: one cannot reconstruct how she is changed or know what elements contribute to that change. Nevertheless, Dover does not provide a rationale for the two properties she attributes to changes of one's self-understanding in interpersonal inquiry. I argue that these two traits are explainable in Gadamerian hermeneutics.

5.1 Unpredictability

Suppose one's understanding is at stage n. If one wants to successfully predict what one's understanding would be like at stage $n + 1$, one should know (1) what utterance from the interlocutor is the trigger of changes and (2) how one's own understanding would change in response to it. However, both conditions are difficult to satisfy—one cannot know when one would change from stage n to stage $n + 1$, nor can one know what stage $n + 1$ would be like.

The first condition is hard to fulfill for two reasons. First, the more obvious one, we are generally incapable of predicting what the interlocutor will say, especially in a conversation where their understanding is changing in response to our words. If one can't know what the interlocutor is going to say, then one can't know how oneself will be changed in response. Secondly, one cannot know which utterance delivered by the interlocutor would trigger changes. As interpretation is governed by *fore-conception of completeness*, revision to one's understanding is triggered when the projected whole cannot cohere with the new meaning that arrives, in other words, when what is said by the interlocutor deviates from one's expectation. However, one cannot anticipate whether a revision would happen because if one can foresee a revision happening, an utterance cannot go against one's expectation; then what the interlocutor says must fit within one's expectation, and no revision would happen. Therefore, knowing that an utterance would cause changes beforehand would cancel the necessary condition for the change to happen, so it is impossible for one to predict when a change would occur.

The second condition fails for a different reason. Note that this does not follow from

the failure of the first condition: one can have a sense of how change will proceed without knowing when it will do so. Suppose, counterfactually, I could know that an utterance would force a revision. I would still not know what revision would come to, because the same utterance would cause different changes depending on which question I heard it answering. The content of changes is determined by the question that orients it. However, the question is not raised by the hearer—in Gadamer's language, it simply "occurs," "arises," or "presents itself" based on which gap of the hearer's understanding is made visible by the interlocutor. One cannot locate the gap herself before it is exposed by an utterance, since it is a part of what determines how things are present to us and thus conceals itself. In this respect, questioning is a "passion rather than action" (Gadamer 360), where "passion" denotes what one undergoes passively rather than pursues actively. As a result, prior to experiencing the change itself, it is also impossible for one to predict the direction in which the change would proceed.

The unpredictability of changes to one's self-understanding is especially meaningful for Dover's abdication condition. If one can predict how one's self-understanding might change, then it opens the door for steering the conversation to achieve the kind of change one wants, which is exactly the retention of interpretive sovereignty that Dover objects to. Therefore, unpredictability is not an accidental property; it is a precondition for the abdication of interpretive sovereignty to be possible.

5.2 Untraceability

That the changes are unpredictable does not entail that they are also untraceable. One might not know a person's destination beforehand but can know how she reached there afterwards by tracing her footprints.

The untraceability of changes to one's self-understanding follows from the opacity of *fore-structure*, which is the situation from which interpretation happens. According to Gadamer, "[t]he illumination of this situation—reflection on effective history—can never be completely achieved. . . . To be historically means that knowledge of oneself can never be complete" (Gadamer 301). Tracing how an utterance changed one's self-understanding requires one to compare one's *fore-structure* before and after. In other words, to trace the changes of one's *fore-structure* between t_1 and t_2, one would need to know, at t_2 or later at t_3, the exact content of one's *fore-structure* at both t_1 and t_2. Gadamer does not consider this possible: "[i]t is impossible to make ourselves aware of a prejudice while it is constantly

operating unnoticed, but only when it is, so to speak, provoked" (298). At t_2, one can know the particular prejudice that the interlocutor's utterance has called into question. But the revision is not confined to that prejudice. Under the influence of the *fore-conception of completeness*, one's *fore-structure* is always expected to maintain inner coherence, so a revision at one site reorganizes regions that were not provoked and thus did not come into view. What one can register at t_2 is therefore only the starting site of the change, not its whole extent, and the remainder of one's *fore-structure* at t_1 and at t_2 continues to operate unnoticed. Nor can the deficit be made up at t_3. The comparison between t_1 and t_2 is itself an act of understanding, carried out from a *fore-structure* at t_3 that is again operative yet unnoticed. Worse, that *fore-structure* at t_3 is partly the product of the very change one is attempting to reconstruct. The instrument of the tracing cannot be a part of what it is supposed to trace, which, according to Gadamer, is "due not to a deficiency in reflection but to the essence of the historical being that we are" (301).

6. Conclusion

Dover develops a phenomenological account to show that the kind of interpersonal inquiry she calls "taking one another seriously" is worth wanting and notes that it tends to change the self-understanding of the interlocutors. However, she does not provide an account for the hermeneutic mechanism behind these changes. My argument fills this gap.

When describing the abdication of interpretive sovereignty, Dover mentions that "a conscious, reflective vetting process" must be suspended. In this way, she presupposes a threefold interpretive architecture where comprehension of an utterance, conscious vetting, and changing one's understanding are treated as three separate steps. I showed that this structure cannot be sustained: evaluating an utterance's fit with one's self-understanding is a part of the comprehension process itself, so the evaluation is neither separable nor able to be suspended voluntarily by deliberation; self-understanding is altered no matter whether the vetting process is upheld, so its suspension is not necessary either. Nevertheless, this does not mean that Dover's ideal is problematic. Dover is wrong about how the ideal could be achieved, not in proposing it.

By drawing on Gadamerian hermeneutics, I explain the interpretive mechanism of how one's self-understanding could be changed during interpersonal inquiry, taking Hannah Upp as an example. I also provide an alternative account of how the abdication of interpretive sovereignty can be achieved and explain why one's self-understanding is more prone to

change when the abdication is fulfilled. Lastly, I use Gadamer's hermeneutics to prove Dover's assertion that changes to one's self-understanding are unpredictable and untraceable in interpersonal inquiry.

Dover commits to an ethics-first approach, where she argues that instead of being guided primarily by the norms derived from metaphysical and epistemological theories, our normative aspirations for human relationships should shape our representational choices of the self. However, "ethics-first" does not license "ethics-only": an ethical ideal can be phenomenologically attractive yet rest on a mistaken assumption of how interpretation works during conversation. I supply the hermeneutic aspect of her ideal and show that it is hermeneutically viable, though not in the way she describes. What abdication asks of us is not the suspension of a reflective vetting process, which is not something one can do, but the foregrounding of one's own prejudices and the direction of the *fore-sight* toward oneself. Hermeneutics does not generate ethical ideals, but it constrains which approaches are feasible, and an ideal whose demand cannot be met is therefore a worse ideal however appealing its phenomenology.

Bibliography

Anderson, John R., and Lael J. Schooler. "Reflections of the Environment in Memory." *Psychological Science*, vol. 2, no. 6, 1991, pp. 396–408, https://doi.org/10.1111/j.1467-9280.1991.tb00174.x.

Carruthers, Peter. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford UP, 2011.

Descartes, René. *Principles of Philosophy*. 1644. *The Philosophical Writings of Descartes*, translated by John Cottingham et al., vol. 1, Cambridge UP, 1985.

Dover, Daniela. "The Conversational Self." *Mind*, vol. 131, no. 521, 2022, pp. 193–230.

Evans, Gareth. *The Varieties of Reference*. Edited by John McDowell, Clarendon Press, 1982.

Gadamer, Hans-Georg. *Truth and Method*. 1960. Translated by Joel Weinsheimer and Donald G. Marshall, 2nd rev. ed., Continuum, 2004.

Graff, Delia. "Shifting Sands: An Interest-Relative Theory of Vagueness." *Philosophical Topics*, vol. 28, no. 1, 2000, pp. 45–81.

Heidegger, Martin. *Being and Time*. Translated by John Macquarrie and Edward Robinson, Blackwell, 1962.

Russell, Francey. "I Want to Know More About You: On Knowing and Acknowledging in *Chinatown*." *Stanley Cavell on Aesthetic Understanding*, edited by Garry L. Hagberg and Constantine Sandis, Springer, 2018.

Talbert, Bonnie M. "Knowing Other People: A Second-Person Framework." *Ratio*, vol. 28, no. 2, 2015, pp. 190–206.

Research Paper

Cinema as the Third Interpreter: Interpretation, Recognition, and Misrecognition in *Gaslight*, *The Philadelphia Story*, and *Little Women*

Ziyue Li
Grade 12

Editors' Introduction

This paper argues that cinema is not a neutral window onto character but an "interpreter" that shapes how viewers understand the people on screen, particularly female characters. Through three films, it examines how relationships with others drive the formation and recognition of character. The thesis is worthwhile, and it is developed most convincingly in the chapter on *Gaslight*, which ties its claims to specific visual and sound choices. The discussion of nonlinear editing in *Little Women* likewise shows how film form can structure the viewer's understanding of a character.

The paper's attention to film form is most sustained in the chapter on *Gaslight*. In the chapters on *The Philadelphia Story* and *Little Women*, the emphasis shifts toward plot and the psychology of Tracy and Jo, so that camera work, editing and sound recede behind questions of character.

1. Introduction: Recognition and the Problem of Interpretation

Women in films are not always ignored or made invisible. Rather, they are scrutinized and described. But before they get the chance to define themselves, women are critiqued by characters or the audiences as to who they are, what do they want, and what they are supposed to become. For instance, women are critiqued as either being selfless or selfish in the context of marriage. Although these critiques are seemingly observations, they become part of the self-discovery process for women. Ultimately, women are defined by audiences as much as they are defined by these critiques.

This creates an interesting question: how can we distinguish between an interpretation that recognizes another person and one controls them? Interpretation itself is not necessarily harmful. Because part of our identity develops through relationships, other people may notice qualities that we have not yet recognized in ourselves.[1] Their understanding can help us reconsider who we are and who we may become, making a more complete picture of ourselves. However, interpretation turns into misrecognition when one person is not allowed to answer, revise, or reject that of the other person. Therefore, the real problem is not whether interpretation should exist, but whether the person being interpreted can still participate in shaping its meaning.

To clarify the problem, I will introduce three related concepts: interpretation, misrecognition, and recognition, and I will define them. Interpretation is making a generalized understanding of the identity and experiences of others, as well as guessing their possible futures. Interpretation can be neutral as there is no right answer when making conclusions, and judging others is a constant in all relationships. Misrecognition occurs when an interpretation denies the person being interpreted the ability to answer, challenge, or revise what is said about her. Misrecognition is most harmful when the subject, the person being interpreted, has no power to contest the interpretation, when everything the subject says is taking away from her and used to argue against her, and when there is an extremely narrow representation of the subject. Recognition, does not mean the truth of who regards someone as they really are. Recognition inverts misrecognition in that it regards the subject as a forming participant of the process. Recognition will always allow the subject to engage as an

[1] S. Cavell, *Cities of Words: Pedagogical Letters on a Register of the Moral Life* (Belknap Press of Harvard University Press, 2004); C. Mackenzie and N. Stoljar, eds., *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Oxford University Press, 2000).

active participant, and can manifest through acceptance of the interpretation, revising the interpretation, and contesting the interpretation.

Cinema contributes to this issue. The audience does not get to view or understand the character's interpretations objectively. The cinema decides for us who/what/when we see and hear certain characters as well as which scenes are important. The cinema also decides the ending of the story. In *Gaslight*, we see the lights dim and hear the sounds that Gregory claims do not exist, so the film supports Paula's experience before she finally ensures her perception. Similarly, a performance, an edit, a change in perspective, etc. can make one understanding of a woman seem more real than others. Cinema is therefore not only showing us how characters perceive her, but how it perceives her as well.

Because of this, cinema can be seen as a third interpreter. The first is the character making a judgment about another character. The second is the character who either accepts, denies, or modifies that judgment. The third is the film itself, and makes the audience believe some judgments more than others.

In *Gaslight*, *The Philadelphia Story*, and *Little Women*,[2] the demand for recognition varies. In *Gaslight*, she must first be treated as if someone could trust her and therefore as if someone would care to listen to her talk about her experiences. In *The Philadelphia Story*, the question is whether someone else's judgment about relationships would help her improve as a person, or would help her fill the gaps in her understanding of herself, as opposed to replacing it. In *Little Women*, the relationships offer a positive example and show a greater extent to which a woman deserves control over the pathways of her own future. Overall, the recognition in these films is the authority to trust one's experiences, the right to modify one's identity through relationships, and the right to participate in directing the pathways of one's future. As a third interpreter of this reality, cinema provides or denies support to these criteria in the name of offering a reason, a framework for judgment, and deciding the degree to which a woman's identity is open or fixed.

2. *Gaslight*: Recognition and the Authority of Experience

Gaslight shows the most obvious and standard form of misrecognition. Gregory does not simply misunderstand Paula. He creates a system in which she is no longer allowed to be

[2] *Gaslight*, directed by G. Cukor (Metro-Goldwyn-Mayer, 1944); *The Philadelphia Story*, directed by G. Cukor (Metro-Goldwyn-Mayer, 1940); *Little Women*, directed by G. Gerwig (Columbia Pictures, Regency Enterprises, and Pascal Pictures, 2019).

regarded as a reliable interpreter of anything, including her own mind.

At first, many of Gregory's actions are small. For example, an object disappears, and he tells Paula that she must have moved it. When she hears footsteps above the ceiling, he acts as if there isn't any sound. The gaslights become dimmer, yet no one around her agrees with her about the change. Gregory repeatedly says that she is forgetful and unstable.

Unfortunately, Paula cannot reject Gregory's interpretation. If she says she remembers something, he would say that her memory was unreliable; if she insists that she heard a sound, this insistence would become evidence of her nervous condition; if she becomes emotionally fragile because no one believes her, she is kept at home because of her "mental problem." In this way, every attempt Paula makes to defend herself is turned into further proof that Gregory is right.

More importantly, Gregory does not only misinterpret Paula's experiences; he removes her authority to interpret experience at all, including her own experience.

This loss of authority gradually changes Paula's identity. At the beginning, she may be uncertain about particular events, but eventually she becomes uncertain about herself. She becomes to rely on others like Gregory to explain everything to her, and she comes incapable of behaving normally in public. When she loses trust in her own sense, she also loses the confidence to act her own. Gregory finally becomes the person who defines Paula's reality and therefore defines her.

The film, however, creates a distinction between Gregory's interpretation and the audience's experience. Gregory denies the changes in the gaslight, but the film allows us to see them. Paula hears movement above her, and the soundtrack allows us to hear it too. Though Gregory controls the context of what happens, the cinematic form allows the audience to gradually understand that Paula's perception is based on reality.

In a way, the film supports Paula before a detective does, and before the other characters start to believe her. So, the film provides Paula with a kind of acknowledgment that her husband will not, and that the bystanders will not. It validates her.

Also important is the last confrontation. Paula does not only learn that she was right, she gets back her ability to make judgments. Ironically, she is the one that Gregory needs help from, since he tries to convince her that she is reasonable enough to help him. While Paula's reaction seems confusing and ironic, it is actually a sign that their relationship has changed in regards to interpretation. She is not begging Gregory to figure out what is real once again. She knows him and the threat, and decides for herself.

Gaslight presents epistemic authority as the first condition of recognition, meaning

Paula has a right to her perceptions and the knowledge that comes with possessing them. Paula is not able to recognize herself until her perceptions are given value. *Gaslight* shows us the first condition of recognition. To be recognized, one's perceptions and judgments must be acknowledged.

However, being believed and recognized is not the end. *The Philadelphia Story* shows a situation that is more complicated. Tracy is not Paula's situation, as she is able to defend and express herself and more importantly, recognize and judge her limitations. However, the situation is not that simple. If someone's Judgment is correct and, while doing so, recognizes the other's limitations, how can that be done without assuming control over the identity being formed?

3. *The Philadelphia Story*: Recognition Through Relational Revision

Unlike Paula, Tracy Lord is not depicted as having one dominating interpreter. Rather, she is shown as having multiple people around her who portray her in different lights.

To the public and to the few reporters who manage to enter her house, Tracy is nothing more than a rich socialite. To George, her fiancé, she is almost the ideal woman, elegant, pure, and above the average. Mike perceives her through the sketch of the heiress, and then, starts to see her in a romantic light. Tracy's former husband, Dexter, sees her as the strong and intelligent woman she is, but thinks too highly of herself and is too proud to recognize the weakness in others. Tracy is also aware of how she perceives herself and how she would like others to perceive her.

So, the movie is not about the conflict of truth and deception. Rather, it is the contest of perceptions. The main interest is not in the question of who has the right perception of Tracy, but in the question of which perceptions of Tracy allow her to participate actively in her own self-realization.

George's description of Tracy as a goddess is, in a way, placing her on a pedestal, and in a way, an act of mistaking her, as he is not allowing her the right to be a work in progress.

A goddess is not allowed to be fallible. A goddess also is not allowed to become confused, wishy washy, or show her vulnerabilities that contradict the goddess image. George therefore places Tracy in an image that is a portrayal of himself, rather than a portrayal of Tracy as a complex person.

Acknowledging misrecognition in these scenarios is beneficial because compliments can be a form of misrecognition just as insults can. Praise George's reliance on idealization is

a way to place restrictions on the number of selves Tracy is able to take.

Dexter is more complex to analyze. Just like George, Dexter has a negative opinion of Tracy, and thinks she is cold and inflexible, and that she even lacks the ability to forgive people for being weak. Challenging her is something he enjoys doing, which seems very similar to what George does in order to confine her in a way and draw the boundary, but in Dexter's case, it's fundamentally different.

Unlike Gregory's interpretation of Paula, Dexter's understanding of Tracy remains open to her response. Tracy can argue with him, reject his judgment, and decide whether any part of it belongs in her own self-understanding. While their exchange is not equal in all ways, it is still an exchange. Dexter's take on Tracy is still grounded in the fact that her opinion matters, and he is not engaging in arguments just to demonstrate that her ability to rationalize is fundamentally lacking. Therefore, his take is somewhat valid.

More crucially, Tracy later acknowledges at least some of what he perceives in her. She does not acquiesce to his portrayal. Her nighttime experiences, especially the moments when her systematically cultivated public persona begins to relax, enable her to perceive herself in new and different ways. She becomes increasingly less sure that strength means the capacity to not need others and that moral seriousness equates to the unrelenting moral judgment of others.

The issue is not with whether Dexter is right, and Tracy is wrong. The issue is with whether his perspective is going to be seen as a part of a change that Tracy is willing to accept as her own.

This perspective of recognition suggests a second relational revision. Other people may help us see parts of ourselves that we have ignored, but their interpretations cannot simply replace our own understanding. We must be able to respond to them, test them against our experience, and decide whether they belong in the person we want to become.

Tracy's transformation is therefore not written by Dexter alone. It develops through her encounters with George, Mike, Dexter, and her own unexpected behavior during the night. These relationships offer competing interpretations of her, but Tracy must still decide which of them she recognizes as part of herself.

Tracy's socially and emotionally unembellished performance at the start of the film exemplifies this relational change process. Tracy appears distanced, and her social role creation is evidence of her desire to maintain her social role of unforgiveness. Tracy is socially fragmented and unforgiving towards any relational disruptions to her social role of being distanced.

In later scenes, particularly once drinking starts, her body appears to lose control. Her voice shifts, movements are looser, and reactions are more impulsive, less inhibited. Describing drunk Tracy as the "real Tracy" is an oversimplification. This, too, would be a singular, absolute interpretation. What is important here is that Tracy comes into contact with a previously inaccessible part of herself, that is, an alternative version that had been previously unaccepted in her self-image.

The portrayal Tracy's performance illustrates that self-identity is not a constant abstract that is concealable. Rather, self-identity can emerge from as yet not fully explored actions and relationships.

The film's construction of Tracy's development also is problematic. The presentation of her increased openness, tolerance, and emotional availability as an enhancement of Tracy's character is valid to an extent, as she is ultimately more positive and more connected to her peers. But the film risks suggesting that Tracy becomes more fully human only when she becomes softer, more forgiving, and more suitable for marriage.

The wedding ending exemplifies this tension. Firstly, Tracy actively chooses Dexter instead of continuing with George. She sees George's idealization of her would mean that she lives as a simple person. Tracy also sees that she has been overly controlling, and therefore, the marriage can mean that both people in the marriage can be comfortable with the fact that they know each other's shortcomings.

At the same time, the wedding becomes the film's final proof that Tracy has grown. Her development is confirmed through her return to marriage, which risks narrowing her transformation into the role of a "better" wife. The film, thus, frees Tracy from the goddess role, but almost immediately, places her in the corrected wife role.

This ambiguity shows the brilliance of *The Philadelphia Story*. There is almost no instance of pure recognition or misrecognition in the film. Instead, the film shows the hurdles to recognition through relationships. There can be a person, or even spouses, who can know and relate to. Nevertheless, the recognition must be valid, and there must be the freedom to challenge and/or alter the recognition for the purpose of expanding one's life.

Recognition is more flexible here than in *Gaslight*. In that film, recognition is almost exclusively predicated on belief. In this instance, the recognition of Tracy and the other voices is more important and, therefore, more flexible.

Even if Tracy engages with her transformation, the movie depicts it with a clear finality. The marriage ending contains the transformation. *Little Women* takes the question one step further. It prompts the question of whether a woman can exist without being

completed. It asks whether a film can show her evolving with relationships and be free to suggest different possibilities.

4. *Little Women*: Recognition and the Openness of the Future

Greta Gerwig's *Little Women* provides Jo with a better model of recognition because, without relying on Jo's one final identity, the model allows Jo to be influenced by other people as well.

Oppositional terms are how most people understand Jo. She is the independent one, not the romantic one. She is the writer, not the wife. She is the ambitious one, not the feminine one. The film, however, forces these oppositional terms to be blurred. Freedom, for Jo, is not the absence of loneliness. The absence of bondage is not the absence of the desire for closeness. The memories and relationships that fuel her writing are the memories and relationships of her family. The Jo character can't be defined by a singular identity.

Marmee is an excellent example of a recognizing character. Marmee doesn't deny Jo's anger, nor does she see it as evidence of Jo's innate evil. And contrary to popular belief, instead of dismissing Jo's anger or leaving her to manage it alone, Marmee shares her own experience of anger.

Marmee's answer is important for two reasons. One, Jo is neither stuck with the identity of "angry girl," nor is it declared what Jo must become. Managing this anger becomes a resource for Jo, not a command.

Good recognition allows for self-definition. This is done by providing language and experience.

Marmee's response highlights how self-knowledge requires others. Jo lacked full understanding of her mother because Jo understood that her mother's exterior calmness conceal struggles. Jo and Marmee have both their struggles expanded by the common understanding that recognition is reciprocal.

Jo's sense of self is the product of a web of relationships. Intricate familial relationships with Amy force Jo to foster a sense of balance with competitive feelings, and the frustration of recognizing the limits of women's choices. The dynamic Beth creates with Jo and her eventual absence completely alters Jo's sense of home and memories. Jo sees the limits of the emotional intimacy Laurie offers and the future he wants. Jo's writing gets tangled with a publisher's commercial interests, and her sisters' choices highlight that multiple paths of self-determination exist for women.

Jo is not just restricted by her relationships. Rather, they provide her with materials to write and to know herself. Escaping influence is not what gives Jo her agency. It is her active participation in the processes of interpretation and transformation of these influences. This is called relational self-authorship. This means Jo participates in creating her self together with others. Relationships, memories, social demands, and evaluations of others are part of the process of telling her self.[3]

This idea requires the use of nonlinear editing in the film. Jo's life is not presented as one long line from childhood to adulthood. Instead, the film jumps back and forth in time. The film repeatedly places the sisters' joyful childhood beside scenes of adult distance and loss. Jo's earlier confidence in independence is therefore viewed together with her later loneliness. Amy's childhood frustration is reinterpreted with her adult understanding of money, marriage, and women's limited economic options.

The editing, therefore, performs a kind of recognition. It allows the audience to understand Jo through several times and relationships. Her present does not erase her past, and her past does not completely determine her future.

The ultimate conclusion of the film takes the idea of "openness" even further than before. In one of the endings, Jo receives a romantic reunion with Professor Bhaer. This follows the pattern of the publisher, who suggests that the heroine always marries. In one of the endings, though, Jo observes the printing, binding, and completing of her book. The film leaves the status of the romantic ending uncertain.

The obfuscation is not a trick. It prevents both the marriage and the writing of the book from becoming definitive. The endings contain both the marriage and the negotiation of the marriage, which is Jo's authorship. Jo is a woman who desires love and has a practical view, and a writer who desires love and has a practical view.

The film does not claim that Jo has a lot of freedom. Jo occurs in a world where a woman's life must end in a marriage. Her authorship therefore develops within social and commercial constraints rather than outside them. The film chooses to show the constraints, instead of showing the marriage as the inevitable conclusion.

This is where the film shows the improvement in Jo's authorship. Marmee, her sisters, Laurie, and the publishing world all shape Jo, but no one person can explain Jo completely. The film shows multiple potential outcomes of Jo's authorship. The film ends without completely closing the question of Jo's authorship.

---

[3] Mackenzie and Stoljar, *Relational Autonomy*.

The differences among the three films has become clearer. Paula must first re-establish the power to trust her intuition. Tracy must learn how to adapt herself to the people around her, winning, without relegating herself to becoming someone else's invention. Jo must learn how to allow herself to be influenced, to an extent, without closing off the other possibilities of the future.

These phases move from voice, to dialogue, to possibility. Yet this progression is not available to the audience independently of form: cinema makes recognition perceptible by deciding what is seen, heard, repeated, withheld, and finally closed. The films do not merely record recognition; their audiovisual organization shapes how recognition is experienced.

5. Cinema as the Third Interpreter

Characters interpret one another, but cinema also interprets the act of interpretation itself.

Cinema is not the only medium that interprets. Literature can privilege a narrator's language, and video games can frame a player's choices. What is distinctive about cinema is the way it fixes and synchronizes visual and auditory evidence for a spectator. The camera selects a field of vision; editing controls sequence and duration; performance gives an interpretation a body and voice; and sound can confirm or contradict what a character says. Unlike prose, which usually asks readers to construct these sensory details through language, and unlike games, which can allow players to intervene in or test alternative outcomes, film ordinarily presents a completed audiovisual sequence that viewers cannot alter. Its "third interpretation" therefore arrives not only as an argument about a character, but as a perceptual arrangement through which that character is encountered.

A movie is allowed to portray multiple perspectives concerning a woman, but they are not given the same level of importance.[4] By choosing whose face remains in frame, which reaction receives a cut, what sound accompanies a judgment, and where the story ends, a film directs viewers to specific conclusions and away from others. The third interpreter is therefore not simply another voice in the story; it is also the formal system that arranges the audience's access to every voice.

Through epistemic alignment, interpretive framing, and narrative closure, *Gaslight*, *The Philadelphia Story*, and *Little Women* show three ways in which cinema becomes this

[4] M. Smith, *Engaging Characters: Fiction, Emotion, and the Cinema*, 2nd ed. (Oxford University Press, 2022).

third interpreter.

5.1 Epistemic Alignment: Whose Experience Counts as Evidence?

Alignment concerns how a film controls viewers' access to a character's actions, knowledge, and subjective experience.

In *Gaslight*, the audience notices the lights dimming and the sounds changing. In the case of Paula, the formal choices of the filmmaking give her experience weight, even when everyone around her denies it. The film doesn't just show us that Gregory is dishonest. It shows us the distance between Paula's reality and the way Gregory interprets her reality.

This is important because the process of recognizing something can start long before someone has actually said, "I believe you." The film extends the belief to Paula through its elements of sound, light, and space.

In *The Philadelphia Story*, the alignment is much less direct. The audience is not placed inside Tracy's point of view. Instead, the audience is placed among the many interpretations of Tracy. Because of this, the film still guides us. We see that George's image of worship for Tracy is becoming a rigid image, while Dexter is offering more witty and confident critique of Tracy. Because of this, Tracy's image of her self-control invites us to see her self-control as an image of the self that is incomplete.

In *Little Women*, the audience gets a sense of Jo's feelings in the present through past Jo via the use of a nonlinear editing style. The audience is able to interpret the deeper meaning of Jo's present feelings of isolation and Jo's less Strong and more caring and tender feelings for her family. Throughout the film Jo's present experiences are held to a standard of sometimes a lifetime encapsulated in a single utterance.

Therefore, in cinema, credibility can be organized perceptually. A film can let viewers see the gaslight dim, hear the footsteps, remain outside a character's field of vision, or return to a memory through editing. Literature can describe comparable evidence, but cinema can synchronize image, sound, performance, and duration so that the audience encounters a character's perception as an event unfolding before them. Recognition in film therefore begins with how the medium distributes perceptual credibility: whose experience is made available as evidence, and whose is kept at a distance.

This perceptual credibility is never impartial. When a film puts us in alignment with one individual, it inevitably excludes us from another. The ethical question is not whether the audience is positioned in a certain way, but whether the film allows for the woman's

experience to be more than an external, passive object.

5.2 Interpretive Framing: Which Judgment Appears Insightful?

Cinema's second power is interpretive framing. All narrative media frame, but film does so through a literal visual frame combined with timing, editing, performance, and sound. A novel can describe a judgment as witty or cruel, and a game may let the player accept or reject it through action; film can make the judgment feel persuasive before the audience consciously evaluates it, for example by pairing a line with a reaction shot, comic timing, music, or a particular performance. Framing therefore does not simply tell us what an interpretation means; it gives that interpretation a tone and force.

Gaslight masterfully crafts Gregory's calm and collected demeanor as evidence of manipulation. In this case, Gregory's calm and collected demeanor seems reasonable and is juxtaposed with Paula's increasing distress. However, as the film progresses, Gregory's calm demeanor becomes more of a sinister force. I believe the film teaches the audience to reinterpret his calm and collected demeanor.

In *The Philadelphia Story*, comedy adds another layer to the situation. George's portrayal of his idealization of Tracy is humorous and, in George's case, a bit stiff and humorless. In juxtaposition to George, Dexter has the ability to portray his opinion quickly and with humor, which makes him more flexible and, in a sense, more intelligent as he is able to see through the true nature of the situation.

This actually represents a serious issue. Just because a film makes an interpretation attractive to its audience, it does not mean that the interpretation is ethically correct. For example, the film may find appreciation in Dexter's wit, and he may understand Tracy better than George, but it should be noted that the film grants Dexter more interpretive authority than George. However, the fact that his interpretation appears more attractive does not automatically make it ethically correct.

The third interpreter, cinema, must be examined and not simply trusted. While film may challenge one dominating interpretation, it may equally validate another.

Little Women considers the interpretive authority in unique ways. Marmee, Amy, Laurie, the publisher, and Jo herself all hold parts of Jo's story, and the film gives each of them a voice. Jo's rejection of Laurie does not mean that her later loneliness is false. The film adds to the relativism that exists in the situation.

Recognition doesn't imply a lack of interpretation. It indicates an ongoing opportunity

for correction. One perspective of a film may present one idea as significant, but a film need not present it as an indubitable conclusion.

5.3 Narrative Closure: Who Does the Story Finally Say She Is?

The ending of a film may be its most concentrated interpretive act. Literature also closes narratives, while games may multiply or defer closure through player choice; film, by contrast, typically fixes the final sequence of images, sounds, performances, and cuts that every viewer receives. That fixed audiovisual ending can retroactively organize what earlier ambiguity is allowed to mean.

Narrative closure does more than bring the action to a halt. It provides the audience with a framework for interpreting what has come beforehand.[5] A final scene of marriage, departure, confession, or victory can transform a temporary role into a lifelong legacy.

In *Gaslight*, closure helps to reframe the audience's interpretation of Gregory. Paula is shown to be a credible witness, and Gregory's control is over. The ending helps to restore what has been taken away from Paula by the story. Closure helps to achieve an ethical goal by helping to make public what has been denied in private.

Closure in *The Philadelphia Story* is much more complex. Tracy's second marriage to Dexter shows that she has changed, and that George's romantic idealization was not sufficient for her to form a real relationship. Tracy's marriage to Dexter also becomes the primary symbol of her development. The ending suggests that her story can be said to be complete with her ability to marry the right man.

In *Little Women*, closure is purposefully rendered abstract. The ending of the love story and the ending of the story of publication exist simultaneously. The audience is left to decide whether one is real and one is fantasy, or whether each represents a different phase of Jo's life. The film may be finished, but the audience's interpretation is not.

The three films therefore use closure differently:

Gaslight uses closure to defeat a false identity.

The Philadelphia Story uses closure to confirm a revised identity.

Little Women uses incomplete closure to preserve several possible identities.

This shows why cinema is more than a container for philosophical inquiries. The form of the ending strongly shapes whether recognition appears as liberation, correction, or openness.

---

[5] D. Bordwell, *Narration in the Fiction Film* (University of Wisconsin Press, 1985).

Three criteria of recognition are displayed in these films.

First is credibility. Does the film view a woman's subjective experience and testimony as legitimate evidence?

Second is revisability. Are judgments of her open to be answered and challenged, or is a relational understanding of her forever in the past?

Third is openness. Is her identity allowed to be fluid, or is her identity fixed as the absolute ending?

Cinematic recognition is the result of a film's willingness to view a woman's experience as credible, to keep the possibility of her identity open, and to allow fluidity in the interpretations of her experience. Misrecognition occurs when film form bestows unquestionable power to a specific interpretation of a woman's experience and thereby fixes her identity once and for all.

Avoiding endings, judgments, or clear positions is not what is being asked of cinema. What is being asked of cinema is to pay attention to how meaning-making structures of cinema allocate power.

6. Conclusion: From Being Interpreted to Participating in Interpretation

Recognition is not the same as correctly guessing who another person is.

When a judgment is factually accurate, it can remain controlling if the subject has no opportunity to respond to it. A judgment can be critical and still be constructive because it can provide someone with the language for a change that she can identify as her own. A relationship can mold a person's identity without infringing on their freedom because complete independence from everyone else is not required for freedom.

Gaslight demonstrates that acknowledgment of a person's identity can begin with the power to trust and express one's experience. Without that authority, Paula is unable to perceive and engage with the world as an equal. *The Philadelphia Story* demonstrates that acknowledgment can include arduous modification through relationships, but the individual must retain the capacity to respond and determine whether the change is hers. Recognition in *Little Women* is the ability to provide for one's own future.

These movies illustrate that film cannot exist separately from the struggle for identity. Each close-up, sound, performance, edit, and ending contributes to the way a character is understood. Film may record an experience that has been overlooked in the narrative. It may also indicate that one judgment is wiser than another. Film may acknowledge one identity

and/or validate multiple identities.

Film's ethical obligation is therefore to not stop interpreting women. That would be unfeasible. Its obligation is to not exclude them from the interpretive praxis.

A woman needs the ability to help shape what the influences around her mean. These influences could be things such as her parents, friends, lovers, society and the stories that influence people. Importantly, recognition means that a person is able to answer.

As the third interpreter, cinema does not give a woman's identity back to her. Rather, it provides her with the power to participate in shaping and revising her identity. Its ethical task is not to stop interpreting her, but to ensure that she is never interpreted out of her own becoming.

Bibliography

Bordwell, D. *Narration in the Fiction Film*. University of Wisconsin Press, 1985.

Cavell, S. *Cities of Words: Pedagogical Letters on a Register of the Moral Life*. Belknap Press of Harvard University Press, 2004.

Cukor, G., dir. *Gaslight*. Metro-Goldwyn-Mayer, 1944.

Cukor, G., dir. *The Philadelphia Story*. Metro-Goldwyn-Mayer, 1940.

Fricker, M. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007.

Gerwig, G., dir. *Little Women*. Columbia Pictures, Regency Enterprises, and Pascal Pictures, 2019.

Mackenzie, C., and N. Stoljar, eds. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. Oxford University Press, 2000.

Smith, M. *Engaging Characters: Fiction, Emotion, and the Cinema*. 2nd ed. Oxford University Press, 2022.

Research Paper

Transpacific Sensory Politics: The Translation of Chinese Revolutionary Symbols and the Resistance Aesthetic of African American Artists (1930s–1970s)

Haoyan Shi
Grade 12

Editors' Introduction

This paper examines how African American artists between the 1930s and the 1970s translated the songs and images of Chinese revolutionary culture into a Black aesthetic of resistance. Through the translingual performances of Paul Robeson, the Black Panther Party's reworking of Chinese revolutionary iconography, and the poetry of Langston Hughes, it traces how these artists claimed the right to be heard and seen. Building on Jacques Rancière's "distribution of the sensible," the paper introduces the concept of "sensory segregation" and proposes a three-stage framework of negotiation, productivity and hybridity. The study is ambitious in scope and moves cross-cultural exchange from the level of ideas to the level of the senses. Rather than treating music, poetry and visual art within their separate disciplines, it reads them together as a single transpacific politics of the senses.

The paper's strength lies in its framework and scope. Its readings of the primary materials draw substantially on prior scholarship, and the theoretical lineage of its negotiation–productivity–hybridity framework, particularly its relation to Rancière and Homi K. Bhabha, is left largely implicit.

Editors' note: For clarity, the citations in this article have been converted to MLA format.

1. Introduction: Transpacific Sensory Politics

History often reveals its urgency through stories of voices that dare to cross boundaries. In mid-20th century New York, African American singer and activist Paul Robeson raised his voice in flawless Chinese to sing the Chinese national anthem "March of the Volunteers." For the six thousand people who witnessed it, Robeson's performance was not simply a cultural exchange or aesthetic novelty; it was a deliberate political intervention that unsettled the racialized order of listening in Cold War America, a moment in which sound itself became a claim to public presence and political solidarity. This essay takes that moment as its entry point to examine how, between the 1930s and the 1970s, African American artists translated the sonic and visual idioms of Chinese revolutionary culture into a distinctive Black resistance aesthetic that sought the right to be heard and the right to be seen.

In the mid-twentieth century, African American artists looked across the Pacific to China and found, in its revolutionary symbols, a language for their own struggle for freedom. Through the resonant songs of Paul Robeson, the visual strategies employed in the Black Panther Party's reproduction of Chinese revolutionary iconography, and the poetry of Langston Hughes, they claimed the right to be heard, invented new modes of visibility, and discerned "*third spaces*" where cross-racial solidarity could take root. In doing so, they challenged the United States' racialized "*distribution of the sensible*," turning art into a transnational political aesthetic that linked Black liberation to global anti-imperial movements. This project of negotiation, productivity, and hybridity left behind not only works of art but also a transpacific legacy of sensory emancipation.

The core of this study lies in repositioning transnational cultural exchange from the level of ideas to the sensory level. Building on Jacques Rancière's notion of the "*distribution of the sensible*," I introduce the concept of sensory segregation to name the U.S. racial regime that governs who may be audible or visible, and how those senses are interpreted and policed. Against this background, mid-20th-century China articulated a contrasting politics of the senses: mass songs, posters, and spectacles that explicitly organized sound and sight as resources for collective mobilization. When Chinese revolutionary iconography circulated across the Pacific, they did not translate passively. African American performers, poets, and visual producers reworked these forms to contest the boundaries of audibility and visibility imposed by U.S. racial orders, producing a set of practices I analyze as a transpacific sensory politics.

To capture the mechanisms through which these cross-cultural translations operated, I propose a three-stage analytical framework (N–P–H model): negotiation, productivity, and hybridity. Negotiation denotes strategic interventions into the auditory order: practices that attempt to convert "noise" into legitimate speech or song, thereby claiming a right to be heard. Productivity refers to the active making of new visual grammars and media forms for visibility, including the formatting, layout, and circulation practices that render formerly invisible subjects legible as political actors. Hybridity names the third spaces in which distinct sensory concepts are braided together to produce new collective perceptions and solidarities. Together, the N–P–H model explains not only the influence of Chinese revolutionary signs but how they were strategically manipulated across different sensory registers to reconfigure the African American political presence.

Empirically, the research paper mobilizes three case studies that exemplify each stage of the model. Paul Robeson's multilingual performances and recording projects embody negotiation in the acoustic register; his use of Chinese song unsettled the listening habits of the Anglophone world and momentarily displaced the sonic color line. The Black Panther Party demonstrates productivity through visual practice. By reworking Maoist poster aesthetics and slogan forms into Panther layouts and street spectacles, the Party produced powerful visible claims on public space and political dignity. Langston Hughes's poems and public interventions illustrate hybridity. His transoceanic imagery forges a poetic third space that links anti-colonial struggles with African American demands, producing affective bonds across distance. These artistic acts are analyzed in tandem with the state countermeasures, such as passport revocations, Counter Intelligence Program (COINTELPRO), and House Un-American Activities Committee (HUAC) interrogations, that sought to reseal the sensible. They reveal a dynamic contest in which the push for new sensory freedom continually met by the state's coercive efforts to contain them, shaping an ongoing struggle over sensory politics.

Methodologically, the study combines archival research, close reading of literary and visual texts, and media history analysis of circulation practices (records, newspapers, journals, posters). The materials reproduce a distinct archive from concert programs and records to newspaper layouts and poems, precisely because the politics I describe traverse media and modalities. Sound, print, and image each functions as a site of contestation over presence and meaning.

This project contributes in three main ways. First, it reconstructs a neglected transpacific circuit of cultural exchange in which Chinese revolutionary aesthetics functioned as resources for African American political imagination. Second, it reframes cross-cultural

borrowing as a processual politics of the senses, articulated through the N–P–H model, rather than as simple influence or imitation. Third, it connects historical practices to contemporary debates about representation and public space by arguing that struggles over who may be heard and seen remain critical to cultural politics. The paper proceeds as follows: section two traces the historical and media circuits that carried Chinese revolutionary symbols to U.S. shores; section three elaborates the theoretical frame; section four analyzes the three case studies; section five examines state backlash and contraction; and the conclusion probes the legacy and contemporary reverberations of transpacific sensory politics.

2. Historical Context: Distribution of the Sensible, Double Sensory Confinement, and the Flow of Transnational Revolutionary Symbols

2.1 Political Aesthetics and "Sensory Segregation"

In *The Politics of Aesthetics*, French philosopher Jacques Rancière introduces the concept of "*distribution of the sensible*," which is not only an aesthetic theory but also a political tool for analysis. For him, art matters politically because it can intervene in and redraw the sensible order: "Artistic practices are 'ways of doing and making' that intervene in the general distribution of ways of doing and making as well as in the relationships they maintain to modes of being and forms of visibility" (Rancière 13).

In this sense, the "*distribution of the sensible*" is the institutional division in society of what can be seen and what can be heard, a division that inherently involves power dynamics about who has the right to speak in public space, whose image is given meaning, and whose existence is systematically ignored. Even when African Americans are physically present, their voices are often muted or distorted and their images reduced to stereotypes, which, as scholar Jennifer Lynn Stoever notes, "makes palpable the visual representational processes that render black people either invisible or hypervisible, but never truly seen and known" (9). This sensory deprivation pushes them to the margins and confines their ways of expression.

Building on these, a theoretical framework of "sensory segregation" is used in this essay to describe the U.S. racial order: a regime that exceeds spatial separation to govern perception itself—controlling who is audible and visible, degrading some voices to "noise," and fixing images through stereotypes. This framework will guide the following analysis of how African American artists translated Chinese revolutionary symbols to reshape a racialized sensory field in the United States.

2.2 Racial "Sensory Segregation" in the United States

2.2.1 The Control of Auditory Order

In the first half of the twentieth century, U.S. cultural industries subjected Black and white music to starkly different regimes of production and circulation. As Stoever observes:

> Frustratingly, African Americans did find themselves listening in as white American broadcasters, networks, advertisers, radio stations, writers, casting agents, musicians' and actors' unions, newspaper critics, and media theorists excised them from contemporary conversations about radio and its role in shaping "modern" citizens. (Stoever 231)

Within Black communities, spirituals, blues, and jazz carried deep political memories and meanings of resistance. Yet when these genres entered white-dominated entertainment markets, their political edge was routinely stripped away, changed into softer, sentimental performances designed for leisure. Black music was thus recast as consumable and non-threatening, tailored to the white audience majority's aesthetic tastes and political safety zones.

Within this structure, sound itself became a key instrument of racial differentiation. As Stoever notes, "the sonic color line enabled the dominant white culture to classify particular sounds as identifiably and essentially 'black'" (36), coding Black voices as emotional and unpredictable rather than neutral public expression.

Stoever further underscores that this process regulated not only sound but also listening itself: "Tethering both an evolving battery of sounds and a limited range of listening practices to black bodies expanded white racism to include new forms of acoustic disciplining," she argues, "that punished racial transgressions and served as violently coercive psychological conformity" (36). Here, "acoustic disciplining" refers to the broader system of auditory control through which the dominant culture trained listeners to hear Black voices in narrowly racialized ways, while punishing sounds that transgressed those boundaries. Even when African Americans were physically present in public space, their voices were often muted, distorted, or reclassified as "noise" or "entertainment."

This form of auditory segregation illustrates a broader logic of "sensory segregation." It curtailed African Americans the full audibility of their cultural expression and enforced psychological conformity through listening practices designed to uphold racial hierarchies. Overcoming such regimes thus requires a redistribution of the sensible, one that makes suppressed voices to be heard as legitimate contributions to the cultural and political

commons.

2.2.2 The Discipline of Visual Order

Alongside auditory control, Black visual representation has long been governed by racialized logics of visibility. Saidiya Hartman urges attention to the everyday scenes where power naturalizes itself: "We do better to cast our glance at the more mundane displays of power and the border where it is difficult to discern domination from recreation. By disassembling the 'benign' scene, we confront the everyday practice of domination, the daily routine of terror, the non-event, as it were" (Hartman 66). These "benign" spectacles such as coerced entertainments, staged dances, and curated pleasures install a visual order that renders domination ordinary.

As the article succinctly puts it, "The Saturday games at plantation house acted as a productive force that regulated the forms of expression allowed to the dominated" (Pereira). In other words, visibility is not merely reflective but productive. It manufactures permissible gestures and affects, dictating the very terms of appearance.

The slave trade revealed this logic in its starkest form. The act of display forced victims to appear as if they were willing participants, a performance meant to render the transaction orderly and respectable. As Hartman notes, "To 'make themselves sellable,' the captives chained by the neck were forced to sing Old Virginia Never Tire . . . whose essential function . . . was to make the subjugated appear content and comfortable" (Hartman 49). In this cruel staging, joy and suffering were placed side by side, with a mask of cheerfulness imposed to obscure the violence beneath.

The crucial question of visual politics, then, is how to dismantle the scenes in which violence remains unseen. Hartman's provocation is blunt: "Is the perfect picture of the crime the one in which the crime goes undetected?" (66). This challenge aligns with the paper's broader account of sensory segregation: even when Black bodies appear in public imagery, they are contained within visual regimes that foreclose speech and deny agency.

2.3 The Sensory Politics of Chinese Revolutionary Symbols and Their Transnational Dissemination

In stark contrast to the racialized sensory segregation in the United States, China developed a sensory politics of revolutionary culture during the mid-20th century. This model used sound and vision as tools for political mobilization while emphasizing direct mass participation.

2.3.1 The Acoustic Power of Revolutionary Music

Unlike the U.S. racial regime of "sensory segregation," 1930s China developed a musical politics centered on mobilization. The "March of the Volunteers" became one emblem. As Howard notes, "'March of the Volunteers' created a new national style characterized by its robust, resonant, heroic, and powerful sound—a fervent, vigorous quality previously unknown to Chinese songs" (136). This unprecedented sound turned music into a political force.

Its power lay in merging local and global traditions. Howard emphasizes, "the success of 'March of the Volunteers' lay in Nie Er's ability to use the Chinese musical foundation while incorporating compositional techniques from foreign revolutionary songs and marches . . . " (Howard and Lau 136). By blending labor chants, folk-like melodies, and Soviet "mass song" rhythms, it resonated throughout China and internationally.

These "mass songs" were crafted as tools of collective action. Howard observes that the National Salvation Song Movement positioned song as its primary vehicle for mobilizing national solidarity, and that "it was Nie's 'mass songs' that formed the canon of socially accepted art" (137). The anthem exemplified a sonic politics that ordinary people could readily adopt and perform.

2.3.2 The Dissemination and Political Significance of Chinese Revolutionary Visual Symbols

If the "March of the Volunteers" embodied sonic mobilization, Chinese revolutionary culture also built a politics of the senses through a visual system. From the 1950s, China projected images of workers, peasants, and soldiers and set against red backdrops and bold slogans through journals like *China Pictorial* and events such as the World Festival of Youth and Students (WFYS). As Wilcox notes, the WFYS "was an important platform for cultural exchange among the global left" (35), enabling Chinese artists to present their works and forge ties with foreign communities. Posters, performances, and photographs not only conveyed anti-imperialist and anti-colonial messages but also modeled a collectivist aesthetic that stood in sharp contrast to Western individualism.

These visual practices were circulated not only on stage but also through widely distributed print. *China Pictorial* regularly published images of dances and collective formations, such as "Red Silk Dance in Moscow," originally published in *China Pictorial* (Renmin huabao) 1952, No. 1 (Wilcox 35). Through images and performances, revolutionary aesthetics became powerful resources for international leftist solidarity.

At the same time, the *Beijing Review* targeted global readers with both text and

images. As Terrell explains, it was "intended for foreign consumption" (193), reaching subscribers in more than 150 countries, from intellectuals to diplomats. This circulation extended revolutionary imagery into transnational spaces, where it made its way to African American radicals seeking alternatives to the racialized visual order of the United States. Chinese revolutionary symbols conveyed political content while simultaneously offering a sensory template for imagining new forms of collective visibility.

2.3.3 The Re-signification of Symbols Across Oceans

Chinese revolutionary symbols acquired new meanings when they entered the United States. For African American artists under racialized "sensory segregation," their mass-oriented and anti-imperialist qualities offered an alternative model for expression beyond the confines of white-dominated culture.

For example, when Paul Robeson sung the "March of the Volunteers" with its robust sonic style noted by Howard (136), the performance transcended the original Chinese context and reanimated the anthem as part of a global repertoire of resistance, aligning its heroic resonance with anti-racist and anti-colonial struggle. This transformation illustrates how revolutionary symbols circulated across borders not as static artifacts but as mutable resources actively re-signified in challenging oppressive sensory regimes.

3. Theoretical Framework: From Sensory Subversion to the Creation of Third Space

In section 2, we observed the stark contrast between the racialized "sensory segregation" in the United State and China's emphasis on the "participatory" nature of its revolutionary culture. However, to fully understand how African American artists translated Chinese revolutionary symbols into their own political aesthetics, we need a theoretical tool that can explain how power structures achieve cross-cultural hybridity and fusion in cultural reproduction. This section will integrate Jacques Rancière's theory of the "*distribution of the sensible*" with Homi K. Bhabha's concept of the "*Third Space*," proposing a three-stage analytical framework: Negotiation–Productivity–Hybridity, abbreviated as the N–P–H model.

3.1 Cross-Cultural Translation and the "Third Space"

For Homi Bhabha, the "*Third Space*" is not a fixed or final cultural location but rather a space where new meanings are created in the process of cultural exchange. Identity and difference are not permanent but are continuously reshaped as people move between different

contexts. Bhabha explains: "we find ourselves in the moment of transit where space and time cross to produce complex figures of difference and identity, past and present, inside and outside, inclusion and exclusion" (2). This means cross-cultural translation emphasizes how symbols are reinterpreted and re-understood in new settings.

The *Third Space* is where cultural differences are not simply given or static but are constantly made and remade through people's actions and interactions. As Bhabha puts it: "Terms of cultural engagement, whether antagonistic or affiliative, are produced performatively. The representation of difference must not be hastily read as the reflection of pre-given ethnic or cultural traits" (Bhabha 3). In this sense, the *Third Space* serves, rather than a compromise between cultures, as a dynamic site where negotiation and production give rise to new identities and political positions.

For African American artists, incorporating Chinese revolutionary symbols meant actively engaging with and transforming them. In the *Third Space*, these symbols were re-narrated, connecting anti-imperial, mass mobilization with the political struggles of the Black community in America. As Bhabha says: "The borderline work of culture demands an encounter with 'newness' that is not part of the continuum of past and present. It creates a sense of the new as an insurgent act of cultural translation" (Bhabha 10). This process of "cultural translation" allowed the revolutionary symbols to break away from their original meanings and speak directly to the challenges of racism and colonialism in the United States.

3.2 "Distribution of the Sensible" and the Redistribution of Sensory Power

In Jacques Rancière's account, the "*distribution of the sensible*" names the perceptual order that decides how a common world appears and who counts within it. In other words, politics is not an abstract idea but about redrawing the boundaries of who is visible, audible, and excluded as silent or as mere noise.

On this basis, art matters politically because it can alter the ownership of voice and visibility. In our case, African American artists' cross-cultural translation of Chinese revolutionary symbols works as a deliberate redistribution of the sensible. Within a U.S. racial order that often renders Black voices inaudible and Black images misread, importing these sonic and visual grammars created new public events that renegotiated the borders of who may speak and who may be seen, precisely where Rancière locates the core of politics.

3.3 The N–P–H Three-Stage Analytical Framework

Combining the concepts of "*Third Space*" and "*distribution of the sensible*," this section introduces the three-stage analytical framework Negotiation–Productivity–Hybridity, to explain the operational mechanism of cross-cultural symbols in the context of Black art.

3.3.1 Negotiation—The Struggle for the "Right to Be Heard"

"Negotiation" refers to the strategic contest over who gets voice heard under unequal power. Rancière points out that the sensible field already defines who can take part and who is excluded, determining which voices and images are recognized or ignored. As he described, "the distribution of the sensible thus produces a system of self-evident facts . . . of what is visible and audible . . . 'distribution' therefore refers both to forms of inclusion and to forms of exclusion" (Rancière 85).

In the United States where Black voices were routinely marginalized, African American artists had to negotiate within the dominant listening order, seizing openings to unsettle it and make themselves audible. As Stoever observes, these racialized sonic politics "profoundly impacted the ability of black people . . . to claim, enact, and sound their rights in American life" (Stoever 34). This highlights how Black artists used strategic acts of expression and cross-cultural performance to negotiate their presence and transform suppressed sound into the right to be heard.

Negotiation, then, aims to reframe voices that have been treated as "noise" into legitimate speech. For example, Robeson's rendition of "March of the Volunteers" in Chinese broke the hegemony of anglophone speaking, introducing the audience to an unfamiliar auditory environment and thereby briefly displacing the existing racialized listening mechanism. This served as an active intervention in the established "hierarchy of sounds," aiming to secure the right to be heard for Black voices.

3.3.2 Productivity—The Struggle for the "Right to Be Seen"

"Productivity" refers to making new visual codes that reorganize public visibility and reconsider who appears and how they appear, necessary because dominant visual regimes have long constrained Black representation.

On the ground, Chinese revolutionary signs moved into Black communities via external publications, delegations, and grassroots circulation, and were then reedited into a new visual grammar. Kelley and Esch note that "In the late 1960s and early 1970s, it seemed as though everyone had a copy of Quotations from Chairman Mao TseTung, better known as

the ‘Little Red Book.’ From time to time supporters of the Black Panther Party would be seen selling the Little Red Book on street corners as a fundraiser for the party” (98). Such street-level insertion, alongside newspapers and posters, produced an alternative visibility: red iconography, slogans, mass images, and poses of armed self-defense were juxtaposed to form a transnational anti-imperialist narrative that confronted audiences with a different political perspective.

This productive work also shifted affective position. Kelley and Esch record Huey P. Newton’s description of his China trip as “a sensation of freedom . . . I felt absolutely free for the first time in my life—completely free among my fellow men,” an experience that Black Panther media translated into public visual testimony (Kelley and Esch 98). As Judson Jeffries observes, the introduction to *Black Against Empire* “begins with an interesting . . . vignette of Huey P. Newton’s visit to the People’s Republic of China in the fall of 1971,” placing the Panthers within a transnational frame of visibility (Jeffries 180). In short, productivity here means not passive exposure but the making of visibility through translation, juxtaposition, and reediting, so that different people, images, and claims can be seen in new ways.

3.3.3 Hybridity—Creating the “Community of the Perceptible”

“Hybridity” demonstrates not a simple mix-and-match of cultures, but a lived process of contact, negotiation, and production through which distinct sensory experiences (what is seen, heard, and felt) are braided into a shared field of perception and action. Langston Hughes sets Chinese anti-imperialist signs, images, and sounds alongside Black radical traditions so that each becomes audible and visible to the other, contracting distant communities into a perceptible common. As Lai-Henderson notes, Hughes “rewrites the public image of African Americans in the Chinese cultural and intellectual imagination,” and his work “link[s] local-regional to outer-national sites . . . [conveying] a kind of universalism that transcends racial, national, and sociopolitical boundaries of the American color line” (Lai-Henderson 104).

Hughes’s 1937 poem “Roar, China!” stages a striking juxtaposition: it places China’s struggle against foreign domination alongside the United States’ Jim Crow order, collapsing distance into a single sensory map of power. The poem calls to “Smash the revolving doors of the Jim Crow Y.M.C.A.s. . . . Stand up and roar, China!” (Gao, ch. 5). This coupling invites readers to perceive Afro-Asian oppression as structurally connected and to experience a shared claim to visibility and audibility across borders.

Hybridity functions as a sensory reprogramming. By adapting Chinese revolutionary signs to Black political aesthetics, Hughes makes each tradition witness and amplify the other, an act through which a cross-racial perceptible community comes into being.

3.4 Summary of the Theoretical Framework

The N–P–H model is not a linear process, but rather a dynamic system in motion. In reality, negotiation, productivity, and hybridity often intertwine, and each stage may face resistance and compression from power structures. However, this model allows us to clarify how the practices of specific artists unfold across different sensory domains and generate new political aesthetics through cross-cultural translation.

In the next sections, this study will analyze Robeson's acoustic politics (Negotiation), the Black Panther Party's reproduction of Chinese revolutionary visual symbols (Productivity), and Hughes's spiritual hybridity (Hybridity), revealing how these practices collectively form a "sensory liberation chain" and leave a lasting political legacy within the transpacific artistic network.

4. Case Studies: Three Practices of Sensory Reconstruction

4.1 Negotiation / Acoustic Politics: Paul Robeson and "March of the Volunteers"

4.1.1 Historical and Political Context

Paul Robeson (1898–1976) was not only one of the most acclaimed bass-baritone singers and actors of the twentieth century but also a figure who constantly fused art with politics. His early successes on Broadway and London's West End soon made him an international star. This visibility gave him a unique platform to bring Black spirituals, workers' songs, and revolutionary anthems to audiences from the Soviet Union to China.

Robeson always treated performance as political. He insisted that artists must stand with the people, not above them. As Duberman cites him saying during the Spanish Civil War: "It is dishonorable to put yourself on a plane above the masses. As an artist I come to you; I am one of you" (Duberman 261). This statement reveals his conviction that art was not neutral but a weapon in anti-fascist and anti-colonial struggle.

With the onset of the Cold War, the U.S. political climate hardened. Under McCarthyism, Robeson's internationalism and ties to the Soviet Union and China made him a prime target. In 1950, the State Department revoked his passport, cutting him off from

international stages. As Shana L. Redmond explains, “The first is his passport revocation . . . [1950–1958] . . . effectively ended the superstar stage of his career, killed his music industry recording opportunities, and isolated him . . . He simply would not be kept from his communities” (Redmond 8).

Duberman confirms that Robeson denounced the action as nothing less than silencing: “Robeson himself denounced the passport action as one more attempt by the Truman administration ‘to silence the protests of the Negro people’—but his statement was not widely carried in the press” (387).

Yet Robeson refused to be muted. As Redmond observes, he would not be kept from his communities. Instead, he invented new techniques of listening and for being heard. In the years of passport revocation and media restriction, he continued to perform domestically, record, and work through peace networks and allies abroad. This persistence itself was a political act—an insistence on reclaiming the right to be heard.

4.1.2 The Cross-Cultural Negotiation of Sound

Amid early Cold War repression and racialized listening in the United States, Paul Robeson chose to carry “March of the Volunteers” (“Chee Lai!”) onto American and international stages, turning a Chinese wartime anthem into a multilingual, cross-border public event. His collaboration with Liu Liangmo began with attentive listening and notation. At New York’s City College stadium, “he announced, ‘I am going to sing a Chinese fighting song, called “Chee Lai!”’ . . . he sang in perfect Chinese; the audience demanded an encore” (Gao, ch. 2).

He then scaled up the song’s reach. At Washington’s Union Square and New York’s Madison Square Garden, Robeson stood “shoulder to shoulder with Liu and performers of the Chinese People’s Chorus,” singing “‘Chee Lai!’ in both Chinese and English” (Gao, ch. 2). To embed the anthem in everyday listening, Robeson, Liu, and the Chorus recorded *Chee Lai! Songs of New China* (Keynote, 1941) with the clear aim that “the record [be] played on radio stations and in homes around the United States” (Gao, ch. 2). The sound then traveled further: “Hollywood eventually adopted ‘Chee Lai!’ . . . The U.S. Army Air Force Orchestra played the tune at the start and conclusion of the film Why We Fight: The Battle of China (1944)” (Gao, ch. 2).

Within this frame, Robeson’s triad of Chinese lyrics, bass timbre, and mass media redistributed audibility. Live performance unsettled Anglophone dominance, stadiums and squares created collective listening spaces, and records, radio, and film extended the reach of

revolutionary sound. Within an unequal political and cultural order, he used cross-linguistic performance and circulation to compel audiences to confront voices from the margins, which greatly exemplified the meaning of negotiation. This fleeting "acoustic displacement" was a direct claim and rewriting of the right to be heard.

This is a clear example of "negotiation" in the N–P–H model: strategically using voice and language to contest the right to be heard in an unequal power structure, displacing the existing racialized auditory order, if only momentarily.

4.2 Productivity / The Black Panther Party's Reproduction of Chinese Revolutionary Visual Symbols

4.2.1 Visual Production: Historical and Political Background

In the 1960s and 70s, the Black Panther Party answered state violence with armed self-defense, but it also treated culture as a central front. The Party's newspaper, *The Black Panther*, was the key instrument because it fused political ideas with a visual language. Huey P. Newton set a clear cultural rule—"the only real culture is revolutionary culture" (Seale)—so images, slogans, and layout worked for mobilization. The paper also sustained the Party materially: "One of the main sources of funds is the Party's newspapers. It is an organ which lumpen proletarian brothers and sisters produce" (Seale). Readers were not passive: "A large portion of our money comes from many groups and people . . . who sympathize with the Party because they have taken time to read our newspaper" (Seale). These lines make "cultural production" concrete: the paper placed Black communities into public visibility and tied the organization to its base.

At the same time, the Panthers brought Chinese revolutionary culture into U.S. streets and turned it into a workable visual material tactic. The strongest example is the "Red Books for Guns" drive. Seale records Huey's idea: "We can make some money to buy some guns . . . We can sell those Red Books" (Seale). They framed it literally as "Dollars for Red Books." At Berkeley's gate, Huey turned text into a public spectacle: "Power comes out of the barrel of a gun. Quotations from Chairman Mao Tsetung. Get your Red Book" (Seale). The outcome was immediate: "We sold the Red Books inside of an hour . . . the next thing I knew, we had enough money to buy two shotguns" (Seale). This method made a small red book into a street icon and translated anti-imperialist discourse into a handheld visual object. It shows how the Party's cultural and revolutionary significance lay in making politics visible: connecting public imagery and resource mobilization, while embedding Chinese elements

within the graphic vocabulary of Black liberation.

4.2.2 Producing the "Right to Be Seen"

In the late 1960s and early 1970s, *The Black Panther* used layout as a visual instrument to produce political visibility. Methodologically, the Party's pages resonated with Mao-era China: both treated images as "conduits of values and political culture" (Gaiter 90). This helps explain the paper's heavy reliance on illustration and poster-like spreads, since, as Huey P. Newton put it, "The black community is basically not a reading community . . . but learns through observation of or participation in [an] activity" (Gaiter 90). The Panthers thus translated Maoist visual grammar (high-contrast black/white/red, slogan-like text, massed figures) into the U.S. context, and circulated their own claims outward; the paper "was shipped to locations all over the world, including China" (Gaiter 96). Emory Douglas's bold line-work often "alluded to prototypical simple black, white, and red Maoist posters," refitting Third World revolutionary directness into a street-readable Black liberation language (Gaiter 96). Together these practices carved out visibility within a racialized visual order.

A distinctive proof is the cover of 16 March 1969: a near full-page portrait of Mao with the caption, "Without a People's Army, the People Have Nothing." The image was black and white, overlaid by a "bright red transparent" field—"Chinese red"—with a collaged "army of young people with rifles raised." The back cover enlarged the same background, the red overlay signifying "Red China," the Cold War enemy of the U.S. but "the friend of international revolutionaries" (Gaiter 96). Rather than mere borrowing, the page recomposed Chinese revolutionary tropes with Panther narratives of self-defense, mobilization, and internationalism. Hence, "The Black Panther cover image communicates solidarity between Chairman Mao and the BPP" (Gaiter 97). In the same vein, Douglas "deliberately presented everyday people as heroes," shifting the scale of heroism from leader-images to ordinary Black men and women—so readers weekly encountered themselves as historical actors (Gaiter 102).

This case demonstrates productivity within the N–P–H framework, which uses media and design to generate a new system of meaning that inserts transpacific signs into a local political scene to contest and expand the right to be seen.

4.3 Hybridity / Spiritual Hybridity: Langston Hughes's Chinese Poetry

4.3.1 The Internationalist Poet

Langston Hughes emerges here not only as a major Black American poet with a global outlook but as a writer who linked struggles across continents through travel, friendship, and art. As he once said,

> For the last 20 years, half writer and half vagabond, I have traveled from here to yonder around and back again, up and down the African coast, through Russia, through Asia, back and forth across America and, in general, from pillar to post. (Hughes, *Big Sea* 20)

His internationalism was deliberate. In an unpublished draft of his second autobiography, he wrote that his ambition was "to visit all the dark countries of the world, all those lands where live Negroes or Asiatics. Being an American Negro who grew up in a country predominantly white, it always intrigued me to look into lands predominantly colored" (Hughes, *Big Sea* 15). Early on, he forged ties with Chinese peers. At Columbia he was "closest to a Chinese fellow from Honolulu, Yee Sing Chun, who took him to Chinatown and passed on a smattering of Mandarin," and at that time "Hughes adopted a Chinese pen name, LangHu (or Langhu)" (*Big Sea* 53).

In 1933 he traveled to China and "would see realities that ran counter to the racialized caricatures widely propagated through the American popular media," an experience that broadened his Black internationalist frame (Lai-Henderson 89). This encounter fed directly into his writing, as seen in poems like "Good Morning, Revolution" (1932), "One More S in the U.S.A." (1934), "Roar, China!" (1937), and "Ballad of Lenin" (1938), which, as Lai-Henderson notes, "reflect his advocacy for the kind of democratic ideals that aimed to connect colonial populations and working classes across the globe" (90). The influence of this experience persisted, since "despite the brevity of Hughes's sojourn, China continued to inform his black internationalist position in the African American literary community until his passing" (Lai-Henderson 91).

4.3.2 The Trans-oceanic Imagery in "Roar, China!"

On August 29, 1937, Langston Hughes wrote "Roar, China!" a passionate call for China to resist Japanese invasion. He urged with great intensity:

> Roar, China!
> Roar, old lion of the East!

. . . . . .
Break the chains of the East,
Little coolie boy!
Break the chains of the East,
Red generals!
Break the chains of the East,
Child slaves in the factories!
Smash the iron gates of the Concessions!
Smash the pious doors of the missionary houses!
Smash the revolving doors of the Jim Crow Y.M.C.A.s.
Crush the enemies of land and bread and freedom!
Stand up and roar, China! (Hughes, "Roar, China!" 198–200)

The most striking artistic device in the poem is montage. Hughes juxtaposes different historical scenes to create a cross-national vision. For example, he links the violence of imperialist powers in China with racial segregation in the United States: "Smash the revolving doors of the Jim Crow Y.M.C.A.s" alongside "Smash the iron gates of the Concessions" (Gao, ch. 5). Through montage, he weaves together China's colonial humiliation and African American oppression into a single revolutionary cry.

The poem's emotional power lies in its universal solidarity. Hughes's poetry "uses the imagined 'Chinese' metonym to globalize his own particular experience as an African American within the United States context and to link it to the then-contemporary experience of a Chinese revolutionary reality" (Lai-Henderson 91). This cultural hybridity makes China and African Americans part of the same historical current, both shouting "Roar!" as one.

"Roar, China!" was not an isolated work. Throughout the 1930s and 1940s, Hughes continued to insert China into his writings. From "Song of the Refugee Road" (1940) to "Consider Me" (1951–52) to "Birmingham Sunday" (1963), China reappeared as a bridge linking Black literature and global revolution (Lai-Henderson 91). As Gao notes, "The feelings Hughes expressed in 'Roar, China!' about China's glories, ancient history, recent humiliation, the rise of the Chinese Communist Party, and the imperative for China to 'roar' at this critical historical moment were taken up by W. E. B. Du Bois in his 1959 poem 'I Sing to China'" (Gao, ch. 5). Thus, "Roar, China!" was not only a response to a moment of war, but also a vital node in the larger currents of Black internationalism across the Atlantic and Pacific.

Hughes's poem is a powerful example of the "hybridity" concept, symbolically linking the struggles of oppressed people across distant lands. It creates a "*third space*" where Black Americans and Chinese revolutionaries could find solidarity and a shared political struggle, demonstrating how cultural hybridity forms new spaces for political imagination

and collective action.

5. Power Backlash and the Struggle of Sensory Politics

Yet the practice of transpacific sensory politics was not a one-dimensional path to victory. As Rancière reminds us, any redistribution of the sensible inevitably triggers a backlash from power structures; and Bhabha's "*Third Space*" is not always a stable refuge, but a space that can contract, assimilate, or even be destroyed under political pressure. This section focuses on the three case studies in section 4, analyzing how these artists encountered power backlash at different stages and how these resistances affected the sensory liberation chain.

5.1 Negotiation Tension: The Denial of the "Right to Be Heard"

Amid Cold War repression and a racialized listening order, Paul Robeson's art–politics repeatedly collided with state power. He was not a symbolic "internationalist" in the abstract but a practitioner who used voice to link anti-imperial struggles abroad with anti-racist struggles at home. In August 1950, the U.S. government took a decisive step—"in response to FBI demands, the State Department voided Paul Robeson's passport and informed customs agents that they were 'to endeavor to prevent his departure from the United States' should he attempt to leave'" (Gao, ch. 2). This was not an isolated administrative act but part of an institutional architecture that curtailed audibility. Robeson kept accepting international invitations to keep his case and his right to be heard. For him it was "'especially bitter in this case' . . . because of his deep concern for the world peace movement and his long-felt 'close kinship with the Chinese people'" (Gao, ch. 2). Chinese coverage made the silencing visible as social fact: "people cannot hear Robeson anymore. Hosts of major concerts pretend that he does not exist, officials lock up city halls upon hints of his planned presence at public gatherings, Hollywood is too scared to touch him, and those who had followed him closely for autographs step aside to give way to the FBI investigators" (Gao, ch. 2).

From the U.S. side, Duberman details the routinized surveillance that fenced off Robeson's public voice, such as "having his mail intercepted, his phone conversations bugged, his public appearances monitored and reported" (ch. 21). In 1955 the passport fight hit another wall: "The application was immediately denied; Robeson was again told he had to sign a 'non-Communist' affidavit before a passport for him could even be considered"

(Duberman, ch. 21). Robeson read this as a targeted gag on Black protest, "one more attempt by the Truman administration 'to silence the protests of the Negro people'" (Duberman, ch. 19). Taken together, these statements show how policy, policing, and media protocols synchronized to strip him of public audibility.

5.2 Productivity Tension: The Denial of the "Right to Be Seen"

In 1968, the FBI broadened COINTELPRO program against "black nationalist" organizations and spelled out long-term aims. One directive was to "prevent militant black nationalist groups and leaders from gaining respectability, by discrediting them to three separate segments of the community," and another was "to prevent the long-range growth of militant black nationalist organizations, especially among youth" (Bloom and Martin 202). As the Panthers' influence grew, a September 1968 internal memo sharpened tactics:

> The extremist [Black Panther Party] of Oakland, California, is rapidly expanding . . . [It] is essential that we not only accelerate our investigations . . . but that we take action under the counterintelligence program to disrupt the group. (Bloom and Martin 203)

The program's logic targeted the Party's ability to be seen and heard in public, aiming "to keep this group isolated from the moderate black and white community which may support it" (Bloom and Martin 211).

On the ground, repression combined disinformation, surveillance, infiltration, and lethal force. As Bloom and Martin summarize: "They fed defamatory stories to the press. They wiretapped Panther offices around the country. They hired dozens of informants to infiltrate Panther chapters" (Bloom and Martin 6). Navid Farnia's review-essay underscores the scale: COINTELPRO "targeted the BPP more than any other organization," and records "indicate that the FBI initiated 233 formal counterintelligence actions against the BPP out of the 295 deployed against Black nationalist groups," all within a wider climate that "successfully isolated the Panthers" and used "brute force and constant pressure" to drive the Party to extinction (173). Read together, these sources show a coordinated effort to fracture alliances, discredit leaders, and suppress the Panthers' cultural-political visibility—precisely the terrain where the Party had been most productive in creating new visual and sonic languages.

5.3 Hybridity Tension: The Fragility of Hybridity

From the late 1940s into the early 1950s, Hughes's internationalist poetry was drawn

into McCarthyite scrutiny. In 1953 he was summoned before HUAC, "a hearing that was televised across the nation, although even the FBI admitted that he had never been a communist" (Gao, ch. 5). His poems, especially those touching on China and the Soviet Union, were interrogated line by line. Under pressure he admitted to a "reorientation" of thought. This encounter pushed him into greater caution and self-censorship. Yet as Lai-Henderson stresses, "McCarthyism had a profound impact on how Hughes articulated his political thoughts thereafter, but it did little to stop him from using the powerful tool of internationalism as he continued to address the crises of racism and racial hegemony at home and abroad" (99–100). In other words, his rhetoric shifted from direct slogans to metaphor and symbol, but the framework of international solidarity remained.

This shift is vivid in his poetry. In "Roar, China!" (1937), he called forth a fire-breathing dragon, aligning China with global anti-fascism. By 1963, in his short poem "China," the image became a caged lion: "and looks out on a narrow world— / But is not blind" (Lai-Henderson 102). The change showed hybridity's contraction, from blunt mobilization into symbolic resistance. As Gao observes, "Hughes's suppressed inspiration drawn from the CCP resurfaced in fury at the brutal racial violence African Americans suffered" (ch. 5). Even under Cold War scrutiny, he continued to link the symbolic power of Chinese revolution with African American struggles. Hybridity did not vanish; it survived in allegory, sustaining a trans-racial community of feeling through more cautious expression.

5.4 Conclusion: The Double Cycle of Sensory Politics

Reviewing the three cases reveals how power persistently narrowed the spaces of being heard, being seen, and being sensed together. Paul Robeson, stripped of his passport and silenced through FBI surveillance, lost his ability to carry his voice across borders, embodying the deprivation of the voice. The Black Panther Party's creative visual production met the full force of COINTELPRO, which smeared, infiltrated, and raided its networks, cutting off its right to be seen in public culture. Langston Hughes, after HUAC hearings, shifted from direct revolutionary slogans to allegory and symbol; under McCarthyism, the hybrid "third space" contracted, showing the vulnerability of cross-cultural commons.

Politics operates not only in institutions and force but also in the ordering of the senses. African American artists, through cultural translation and hybrid aesthetics, sought to redistribute the sensible, creating new events of sound and image, while the state worked to reseal those cracks through surveillance, repression, and stigmatization, simultaneously

attesting the power of such sensory politics.

Yet even under pressure, these practices left open an alternative common. They did not overturn the system but preserved other ways of being audible and visible for marginalized groups. This residual sensory community became a resource for imagining transnational solidarity and future cultural politics.

6. Conclusion

This study has shown that the transpacific circulation of Chinese revolutionary imagery and sounds between the 1930s and the 1970s was never a matter of simple ideological importation. Rather, it functioned as a set of sensory resources that African American artists actively translated into practices aimed at remaking who could be heard and who could be seen. At the center of this analysis is the N–P–H model—Negotiation, Productivity, Hybridity—which frames how artists intervened in the racialized "*distribution of the sensible*" to claim auditory and visual presence, produce new public grammars, and construct third spaces that braided distinct political vocabularies. This framework synthesizes the study's theoretical and empirical work and foregrounds the politics of sensory rights as an axis for understanding cross-cultural creativity and repression.

Taken together, the three case studies reveal a recurrent dynamic: creative efforts to expand new sensory possibilities and institutional attempts to restrict them continually collided, creating an ongoing struggle over public perception. Paul Robeson's multilingual performances converted linguistic and sonic "noise" into a political address that momentarily unsettled American listening regimes, but constrained by passport suspensions and surveillance. The Black Panther Party translated Maoist visual vocabularies into street-readable formats that made political dignity legible, but targeted by COINTELPRO and media. Langston Hughes reworked transoceanic poetic imaginations into a hybrid rhetoric of solidarity that survived under the pressures of HUAC and cultural containment. Each case thus demonstrates how the struggle for the right to be heard, the right to be seen, and the ability to imagine a shared public sense was pursued through distinct sensory registers and repeatedly contested by state and institutional power.

This legacy is not confined to the past. It continues to play out in new forms today. The contemporary resonance of these historical dynamics is made plain by the Amy Sherald episode. In July 2025, Sherald announced she would cancel her planned National Portrait Gallery show after disputes arose over her portrait *Trans Forming Liberty* (2024), a depiction

of a trans woman that the artist said the museum had expressed concern about including. According to Sherald, institutional fear, shaped by a hostile political climate toward trans lives, played a role in discussions that might have led to removal. Sherald also reported that the Smithsonian planned to present a video of people reacting to the painting, a curatorial move she opposed because it would have reframed the work as an object for debate rather than as an unmediated claim to presence (Blair). That the dispute unfolded in the shadow of recent institutional pressures further highlights how external political climates and institutional pressure constrain museums' capacities to uphold artists' claims to visibility.

This event mirrors the struggles over sensory space that African American artists fought in the mid-20th century. Sherald's decision should be read as a present-day instance of the "*distribution of the sensible*" I trace historically: efforts to expand who may appear in public and how they appear are continually met by institutional anxieties and political pressures that seek to contain or reframe those appearances. It parallels Robeson's acoustic politics, as Sherald's painting challenges the exclusionary mechanisms of visual order; it echoes the Black Panther Party's visual production, as it too sought to create new "ways of being seen" for marginalized groups through symbolic imagery and cross-cultural elements but faced institutional limits; and it resonates with Hughes's poetic vision, as her artwork attempts to connect the shared experiences of marginalized groups through symbolism and imagery.

From Hughes's poetic imaginations to the Black Panther Party's transformation of Chinese revolutionary imagery into a political language of Black liberation, and Robeson's voice breaking through Cold War auditory blockades, these artists collectively expanded Black presence in the public sensory space. Sherald's exhibition withdrawal reminds us that public spaces must protect the rights to be seen and heard, especially when those rights belong to groups whose sensory presence is most often erased by the sensory order.

The transpacific politics of the senses that influenced mid-century African American engagements with Chinese revolutionary forms were never merely aesthetic experiments; they were insurgent claims on public life, acts of reconfiguring who could be heard, who could be seen, and what could be imagined. Such claims had to be remade continually under shifting regimes of censorship, surveillance, and containment. Yet they carved out fragile openings where new forms of solidarity could be sensed and enacted.

To honor that legacy today is not only to recognize its historical significance but also to insist that cultural institutions and publics confront the political stakes of visibility and audibility. Preserving these conditions means defending the possibility of creating "*third*

spaces," where suppressed voices can resonate, contested images can circulate, and the work of sensory liberation persists as an unfinished yet enduring project within contemporary cultural and political life.

Bibliography

Bhabha, Homi K. *The Location of Culture*. Routledge, 2004.

Blair, Elizabeth. "Artist Amy Sherald Has Canceled Her Upcoming Show at the Smithsonian." *NPR*, 24 July 2025, https://www.npr.org/2025/07/24/nx-s1-5479078/artist-amy-sherald-has-canceled-her-upcoming-show-at-the-smithsonian.

Bloom, Joshua, and Waldo E. Martin. *Black Against Empire: The History and Politics of the Black Panther Party*. U of California P, 2013.

Duberman, Martin B. *Paul Robeson*. New Press, 1989.

Farnia, Navid. "State Repression and the Black Panther Party: Analyzing Joshua Bloom and Waldo E. Martin's *Black Against Empire*." *Journal of African American Studies*, vol. 21, no. 1, 2017, pp. 172–79, https://doi.org/10.1007/s12111-017-9343-y.

Gaiter, Colette. "The Black Panther Newspaper and Revolutionary Aesthetics." *Art, Global Maoism and the Chinese Cultural Revolution*, edited by Jacopo Galimberti et al., Manchester UP, 2019.

Gao, Yunxiang. *Arise, Africa! Roar, China! Black and Chinese Citizens of the World in the Twentieth Century*. U of North Carolina P, 2021.

Hartman, Saidiya V. *Scenes of Subjection: Terror, Slavery, and Self-Making in Nineteenth-Century America*. Rev. and updated ed., W. W. Norton, 2022.

Howard, Joshua H., and Frederick Lau. *Composing for the Revolution: Nie Er and China's Sonic Nationalism*. U of Hawaii P, 2020.

Hughes, Langston. "Roar, China!" *The Collected Poems of Langston Hughes*, edited by Arnold Rampersad and David Roessel, Alfred A. Knopf, 1994, pp. 198–200.

Hughes, Langston. *The Big Sea: An Autobiography*. 1940. Hill and Wang, 1993.

Jeffries, Judson L. Review of *Black Against Empire: The History and Politics of the Black Panther Party*, by Joshua Bloom and Waldo E. Martin Jr. *Spectrum: A Journal on Black Men*, vol. 5, no. 1, 2016, pp. 177–84, https://doi.org/10.2979/spectrum.5.1.12.

Kelley, Robin D. G., and Betsy Esch. "Black Like Mao: Red China and Black Revolution." *Afro Asia: Revolutionary Political and Cultural Connections Between African Americans and Asian Americans*, edited by Fred Ho and Bill V. Mullen, Duke UP, 2008.

Lai-Henderson, Selina. "Color Around the Globe: Langston Hughes and Black Internationalism in China." *MELUS*, vol. 45, no. 2, 2020, pp. 88–107.

Pereira, Allan Kardec da Silva. "De-Romanticizing the History of Slavery in the United States: The Tragic Mode in Saidiya Hartman's *Scenes of Subjection*." *Varia Historia*, vol. 40, 2024, e24025, https://doi.org/10.1590/0104-87752024v40e24025.

Rancière, Jacques. *The Politics of Aesthetics: The Distribution of the Sensible*. Translated by Gabriel Rockhill, Continuum, 2004.

Redmond, Shana L. *Everything Man: The Form and Function of Paul Robeson*. Duke UP, 2020.

Seale, Bobby. *Seize the Time*. Hutchinson, 1970.

Stoever, Jennifer Lynn. *The Sonic Color Line: Race and the Cultural Politics of Listening*. New York UP, 2016.

Terrell, Robert L. "The First 25 Years of the *Beijing Review*, an Official Propaganda Organ of the Communist Party of the People's Republic of China." *Gazette*, vol. 37, no. 3, 1986, pp. 191–219.

Wilcox, Emily. "When Folk Dance Was Radical: Cold War Yangge, World Youth Festivals, and Overseas Chinese Leftist Culture in the 1950s and 1960s." *China Perspectives*, no. 1, 2020, pp. 33–42.

Research Paper

The Bamboo Stove and the Poetry Society: Social Networks and Political Contexts Behind the Literary Activities of the Wuxi Gentry in the Mid-Ming Period

Yiyao Feng
Grade 12

Editors' Introduction

xxxxx

1. Introduction

During the Chenghua reign, two cultural undertakings arose within the local gentry circles of Wuxi 無錫 that would be remembered for centuries. During his home-return of mourning in 1477, Qin Kui 秦夔 (1433–1495), the prefect of Wuchang, helped Huishan Temple recover a bamboo tea stove that had drifted into private hands. This recovery of the bamboo tea stove set off a chain of poetic exchanges. Qin Kui first invited local scholars of Wuxi and then his official friends in Beijing to match his verses, producing two scrolls of poems. Later in 1483, the Wuxi official Sheng Yong 盛顒 (1418–1492), together with Wu Kuan 吳寬 (1435–1504) and a company of Hanlin officials, again exchanged verses on the bamboo tea stove, forming a third scroll. Together the three scrolls are known as the *New Poems on the Bamboo Stove* 竹爐新詠. The second undertaking came in 1482. At the initiative of Qin Kui's father Qin Xu 秦旭, ten elderly commoners of Wuxi formed the Bishan Poetry Society 碧山吟社, built a hall and a pavilion beside the Second Spring of Mount Hui, held elegant gatherings and matched poems, and invited Shen Zhou 沈周 (1423–1509) to paint the *Bishan Poetry Society*.

What prompted these two local cultural events, and what were they meant to accomplish? Were the Wuxi gentry simply adding cultural capital to themselves and to their locality, or did the events carry deeper implications and symbolic weight? In the study of *Elegant Gathering in the Apricot Garden*, Yin Ji'nan found that the elegant gatherings of Ming officials, outwardly pastimes of leisure hours and nominally occasions of literary fellowship, in fact harbored an undercurrent of shared political expression.[1] His finding indicates us that the gathering pictures and verse exchanges of the mid-Ming cannot be understood on the level of refined amusement alone. Did the cultural activities of these Wuxi gentry, conducted locally and in coordination with officials in the capital, hold a comparable depth of such intent?

Studies of Ming literati societies and elegant gatherings, for their part, have mostly taken the official circles of the capital or of Suzhou as their subject, and have rarely examined

[1] Yin Ji'nan 尹吉男, "Zhengzhi haishi yule: Xingyuan yaji he Xingyuan yaji tu xin jie" 政治还是娱乐：杏园雅集和《杏园雅集图》新解 [Politics or amusement: The Apricot Garden gathering and its painting, a new interpretation], *Gugong bowuyuan yuankan* 故宫博物院院刊 [Palace Museum Journal], no. 1 (2016): 6–39.

in depth how local gentry used a single object and a single celebrated site to organize cultural activity across regions. This article takes as its center the three scrolls of the *New Poems on the Bamboo Stove* and the founding of the Bishan Poetry Society. It first reconstructs the making, the loss, and the recovery of the bamboo tea stove, bringing to light the process by which its narrative was reconstructed. It then analyzes how Wang Fu's image of "lofty integrity" took shape in scholarly circles and was deliberately put to use. Drawing next on the official careers of Qin Kui and Sheng Yong and on the politics of the Chenghua court, it considers in turn what each man intended in promoting the verse exchanges. Finally it returns to the Bishan Poetry Society, to explain how a society of commoners could draw in the gentry of Jiangnan, and how its membership overlapped inwardly with the network of participants in the *New Poems on the Bamboo Stove*. On this basis the article goes on to examine the symbolic meaning, within the political context of the day, of the Society of the Nine Elders that the bamboo stove poems and the Bishan Poetry Society took as their model.

2. The Narrative of the Bamboo Stove and the Persona of Wang Fu

The value of the bamboo tea stove was never fixed at the moment of its making. The discrepancies among the surviving accounts themselves reveal how, in the retelling of later generations, it gradually became a relic of Wang Fu 王紱 (1362–1416) and a piece of Mount Hui lore.

Wang Da's *Record of the Bamboo Stove* was written in 1395. According to the essay, an unknown bamboo craftsman made the stove for Monk Xinghai 性海 during he was serving as the abbot of Huishan Temple. Wang Da described the form of the stove in detail, calling it "round above and square below, patterned on the form of heaven and earth," can serve "lofty men and untrammeled scholars." It demonstrates that "no distinction need be made between the noble and the mean." The essay's center of gravity lies with Xinghai, the bamboo craftsman, and the lesson of cultivation embodied in the object; it says nothing of Wang Fu taking part in its making. The related *Preface to Pure Poems on the Bamboo Stove* declares that "a thing cannot make itself precious; it is made precious through a person. A name cannot spread of itself; it is spread through intent." What it esteemed was still the way of Xinghai, and it drew comparisons with the Lotus Society of Mount Lu and with Zhidun's fondness for cranes.[2] These two early texts had already established for the bamboo stove a

[2] Wang Da 王達, "Zhu lu ji" 竹爐記 and "Zhu lu qingyong xu" 竹爐清詠序, in *Zhu lu tuyong* 竹爐圖

mode of explanation in which things are "made precious through a person," yet they did not identify that person as Wang Fu.

In 1488, Qin Kui rearranged the stove's pedigree after he helped the temple find the stove. In his *Record of the Recovery of the Bamboo Tea Stove,* he first cited the poem of "the former worthy of our district, Drafter Wang," and then Wang Da's preface, saying that the stove's fame had grown on these accounts. As for how the stove had left the monastery, he said only that after Xinghai's death the stove too was cast adrift among men. A record written by Lu Jian in the same period states that the stove was made in the early Hongwu years, that it later fell to a great family within the city, and soon changed owners twice, and it likewise passes over the households in question.[3] These accounts write a story of loss and recovery as the return of a famous relic, without acknowledging that the original stove had never been an owner-less castoff.

In 1518, Shao Bao's 邵寶 (1460–1527) *Account of the Bamboo Tea Stove* supplied another line of transmission. As he had heard it, when Xinghai became abbot of Huqiu 虎丘 in the early Yongle years, he left the stove with Pan Kecheng 潘克誠 as a parting gift. The stove remained with the Pan family for more than sixty years. During the Chenghua period, a descendant of the Pans gave it to Yang Mengxian 楊孟賢, and only three years after Yang's death did Qin Kui retrieve it from Yang's elder brother Yang Mengjing 楊孟敬. Shao Bao took care to note that the news came from the husband of his maternal aunt, giving this piece of retrospection the character of a family report circulating within a marriage network.[4] Seen in this light, the stove's supposed drifting among men was in fact a passage from the monastery into the households of a local physician and a wealthy family, from which Qin Kui returned it to the Huishan temple. Qin Kui's silence about the transfer between the Pan and Yang families need not have been deliberate concealment. Even so, the omission rewrote a complicated history of private possession into a clear history of a local treasure restored.

詠, comp. Wu Yue 吳鉞, vol. 4 of *Xishan xianzhe congkan* 錫山先哲叢刊 (Zhonghua shuju, 1922), heng 亨 collection, 1a–4b; see also Tan Xiu 談修, comp., *Huishan gujin kao* 惠山古今考 (Wanli-era woodblock ed.), juan 3.

[3] Qin Kui 秦夔, "Fu zhu chalu ji" 復竹茶爐記, in *Huishan gujin kao*, juan 3; Lu Jian 陸簡, "Fu zhu chalu ji," in the same juan.

[4] Shao Bao 邵寶, "Xu zhu chalu" 序竹茶爐, in *Rongchuntang ji* 容春堂集, juan 2, 1a–b. Elsewhere in the manuscript tradition the title is written with xu 敘; this article follows the juan heading of *Rongchuntang ji* in writing xu 序.

Shao Bao's account also moved the date of the stove's making to the fourth year of Jianwen (1402) and added episodes in which Wang Fu convalesced at the Listening to the Pines Hermitage. Wang Fu painted *Mount Lu* for Xinghai, and received a visit from Pan Kecheng, who came to see the painting. On this telling, the stove was made by a Huzhou bamboo craftsman at the bidding of Xinghai and Pan Kecheng, with Wang Fu inscribing a poem on the spot. Compared with Wang Da's record, the date, the persons, and the order of precedence had all changed. The two versions cannot be forced into agreement. Wang Da's earlier record proves that a bamboo stove existed by 1395. Shao Bao's version reflects how later local scholars understood the stove's relation to Wang Fu, to the Pan family, and to the hermitage.

This narrative shift had already appeared in the poetry exchanges in Beijing in 1483. Sheng Yong wrote that Wang Fu and Xinghai made the bamboo stove for boiling tea, and Wu Kuan, recording Sheng Yong's words, stated that the stove came from Drafter Wang Fu himself.[5] From an inscriber among the early poem scrolls, Wang Fu gradually became the stove's maker and even its owner. The later accounts were no simple errors of transmission. What concerned Qin Kui, Sheng Yong, and Shao Bao was no longer who had designed the stove, so much as what sort of person could confer lasting fame upon an object. The formula "made precious through a person" that Wang Da had proposed found, between the Chenghua and Zhengde years, its most fitting candidate among local forebears.

2.1 Wang Fu's Image and the Character Ascribed to the Stove

That Wang Fu displaced Xinghai at the center of the stove's story has much to do with the formation of his posthumous image. When Wang Fu died in 1416, Wang Hong's *Brief Biography of Wang Mengduan* described him as "lofty and untrammeled in spirit from his youth," possessed of "the air of a true poet in his detachment," and dwelt on his refusal to paint readily even for the powerful and wealthy. Wang Hong summed him up in the line "his resolve could hardly be moved by influence or gain," a judgment still grounded in an eyewitness's assessment of Wang Fu's temperament and his principles of association. Zhang Bingru's *Account of Conduct of the Late Secretariat Drafter Wang Mengduan* went further, saying that he "loved forthright words" and "argued to the utmost with a stern countenance, undaunted by the powerful," and recording that whenever he spoke of climbing Mount Hui and boiling tea in the monks' quarters, "his spirit would suddenly kindle," while he remained

[5] Inscriptions by Sheng Yong 盛顒 and Wu Kuan 吳寬, in *Zhu lu tuyong*, yuan 元 collection, 1a–2a.

indifferent to fame and profit.[6] An upright and unbending character and the life setting of brewing spring water on Mount Hui had entered his biography together, at the latest by the time of his death.

The Wang Fu of the early records was nonetheless no recluse wholly removed from social obligation. Hu Guang's 胡廣 (1370–1418) *Funerary Inscription for Wang Mengduan* says that he painted only when the mood arose, so that a man might seek a work for years without success, yet once inspiration came he could wield the brush in a continuous surge. The poems and prose of Hu Guang, Yang Rong 楊榮 (1372–1440), Jin Youzi 金幼孜 (1368–1432), and others further show that during his service in the capital as a Secretariat Drafter, Wang Fu did paint for Hanlin grandees of rather high rank.[7] This does nothing to diminish his unbending integrity. It tells us that his refusals were aimed at requests that ran against his will; official intercourse as such he did not shun. That Wang Fu's image grew ever more distinct owes precisely to the way biographies and anecdotes drew out of a complicated web of associations a single main thread, his refusal to be swayed by influence.

The story of "hearing a flute by moonlight" recorded in Ye Sheng's 葉盛 (1420–1474) *Diary of the East Waters* (Shuidong riji 水東日記) was the key to the fixing of this image. Hearing the sound of a flute from a neighboring house at night, Wang Fu painted bamboo and rocks in the transport of the moment and gave the painting to the flutist. Learning that the man was a merchant, he did not ask it back, yet when the man came with lavish gifts to request a companion piece, Wang Fu rebuked him angrily, saying, "How could a vulgar fellow be worthy of my brush," retrieved the earlier painting, and tore it up. Ye Sheng called him "a scholar of lofty integrity who rose utterly above the vulgar."[8] By placing the giving and the refusal of paintings in opposite circumstances, the story insists on the conditions

---

[6] Wang Hong 王洪, "Wang Mengduan xiaozhuan" 王孟端小傳, appendix to *Wang sheren shiji* 王舍人詩集, in *Xishan xianzhe congkan* 錫山先哲叢刊, vol. 2 (Zhonghua shuju, 1922), 319–20; Zhang Bingru 章昞如, "Gu Zhongshu sheren Wang Mengduan gong xingzhuang" 故中書舍人王孟端公行狀, appendix to the same work, 321–23.

[7] Hu Guang 胡廣, "Zhengshilang Zhongshu sheren Wang Mengduan mubiao" 徵事郎中書舍人王孟端墓表, appendix to *Wang sheren shiji*, 326–27; Hu Guang, "Wei Yang shuzi ti Wang Zhongshu Mengduan xie zhu" 為楊庶子題王中書孟端寫竹 and "Ti Wang Zhongshu Mengduan wei Yang shuzi hua zhuzhang ge" 題王中書孟端為楊庶子畫竹障歌, in *Hu Wenmu gong wenji* 胡文穆公文集, juan 3 and juan 4.

[8] Ye Sheng 葉盛, "Wang Mengduan yishi" 王孟端遺事, in *Shuidong riji* 水東日記, juan 1 (Zhonghua shuju, 1980), 10.

under which painting and calligraphy may change hands, namely kindred spirit and shared intent. Ye Sheng died in 1474. It shows that the story had entered scholarly memory before Qin Kui recovered the stove. In choosing Wang Fu, Qin Kui, Sheng Yong and his circle were availing themselves of a persona already recognizable in their day.

The difference between Wang Fu and Ni Zan 倪瓚 (1301–1374) lies here as well. Both were former worthies of Wuxi, both famed as painters, and later generations characterized each with the word *jie* 介, unyielding. Ni Zan's *juanjie* 狷介, his aloof fastidiousness, is bound up with flight from the disorders of the late Yuan and with a determination to keep himself unsullied. Wang Fu, by contrast, lived through exile, pardon, reclusion, renewed summons, and attendance on the emperor's northern campaigns, and finally died in official quarters. His fame never rested on a lifelong refusal of office. What earned him his reputation was the steely uprightness, the refusal to drift with the vulgar, that he kept amid passages in and out of office and reclusion.[9] For serving or retired officials such as Qin Kui and Sheng Yong, this made Wang Fu easier to take as a model than any pure recluse of the mountains and forests. He had the aspiration toward the forests and springs of Mount Hui, and also the experience of standing in the ranks of the court. He enjoyed the favor of recognition, yet never used his art to ingratiate himself with the powerful. The purity of the bamboo stove, its endurance of fire, and its use in boiling spring water accorded exactly with this ideal of remaining steadfast while holding office.

Qin Kui's part was to resettle Wang Fu's poetic fame at the Listening to the Pines Hermitage. What moved him to seek the stove was Wang Fu's couplet, "Its vapor steams with the third month's spring rains of Yangxian; its voice carries the autumn on both banks of the Xiang." Having found the stove, he further entrusted the scrolls of both the local and the capital exchanges to the monks for safekeeping.[10] What he restored was thus more than the original object. He restored the bonds among object, poem, person, and place. The stove left the Pan and Yang households and returned to the hermitage. Wang Fu's poem left the old scroll and entered a new round of exchanges. Mount Hui acquired, on the strength of a "former worthy of the district," a piece of lore that could be written onward.

---

[9] Zhang Bingru, "Gu Zhongshu sheren Wang Mengduan gong xingzhuang," 321–23. The present summary of Wang Fu's movements between office and reclusion relies solely on the early biographies; the story of his refusing the Duke of Qian's request for a painting, which accrued layer upon layer from the late Ming onward, is not treated here as a fact of his own time.

[10] Qin Kui, "Fu zhu chalu ji."

3. The Circulation of the Poem Scrolls, Sheng Yong's "Bitter Integrity," and the Network of Two Places

That the recovery of the stove became more than a piece of local antiquarian lore turned on the circulation of the poem scrolls. Between the winter of 1477 and the spring of 1478, two rounds of exchanges took shape around the old stove. The first unfolded in Wuxi, with the participation of Gao Zhi, Lu Mian, and others; both men later joined the Bishan Poetry Society. The second was initiated when Qin Kui went to the capital for his performance evaluation. In his colophon he relates that he first recited his own poems to his fellow townsman, a certain Chen, after which dignitaries at court such as Lu Jian 陸簡, Cheng Minzheng 程敏政, and Zhang Bi 張弼 successively honored him with matching verses, so that "within ten days I had received some forty poems in all." When the scroll was complete, Qin Kui asked Lu Jian for a record and prepared to deposit it at the Listening to the Pines Hermitage on his return south, "to pass down as a future story of the mountains."[11]

This inscription shows that the scroll was no on-the-spot record of a single gathering. There was no need to carry the old object to the capital. Qin Kui had only to recount the recovery and pass round the earlier poems to gather several dozen matching pieces within ten days. The participants included Hanlin lecturers, ministry officials, and supervising secretaries, as well as scholars who had yet to pass the examinations; some were fellow townsmen, others bound by the ties of the same examination year, common office, or literary friendship. The bamboo stove served here as a readily transmissible topic. Its pedigree lay in Wang Fu, the Hui spring, and the Listening to the Pines Hermitage; its poetic imagery lay in purity, in forests and springs, and in the brewing of tea; and it traveled along the lines of Qin Kui's personal connections. Once complete, the scroll returned to Mount Hui, and the inscriptions of famous men of the capital became part of a local site. What the stove connected was exactly this reciprocating process, produced in the locality, given weight in the capital, and brought back to the locality for safekeeping.

The poetry exchanges of 1483 followed a similar route, though with changes in the organizer and in the center of meaning. In 1479, Wu Kuan visited the Listening to the Pines Hermitage at Mount Hui, saw the old stove, and drew water from the Second Spring to brew tea. Several years later, Sheng Yong's nephew Sheng Yu made two replica stoves and carried them to the capital, one presented to his uncle and one to Wu Kuan. Wu Kuan led off, and

[11] Colophon by Qin Kui, in *Zhu lu tuyong*, zhen 貞 collection, 13a–14a.

Cheng Minzheng, Ni Yue 倪岳, Sheng Yong, Xie Qian 謝遷, Chen Jue 陳璚, Lu Jian, and others matched his verses in turn. Li Dongyang 李東陽 also wrote a poem, *Matching Sheng Shunchen's Newly Made Bamboo Tea Stove*, which survives through Shao Bao's *Record of Mount Hui* although it was never copied into the extant scroll.[12] Unlike Qin Kui's recovery of the original stove, this round required no proof of the object's antiquity. The value of the replicas came from their form, their naming, and the friendships around them. The bamboo stove thus passed from an old possession of the Listening to the Pines Hermitage into a literati token that could travel with its owner and be made anew.

The makeup of this group of participants is telling. Wu Kuan, Li Dongyang, Cheng Minzheng, Ni Yue, Xie Qian, and most of the others held posts in the Hanlin Academy or in the lecture establishment of the heir apparent, while Chen Qiong was a supervising secretary. Sheng Yong, the senior of them all, had once served as a censor and had just been recalled from the post of left provincial administration commissioner of Shaanxi to become right vice minister of the Ministry of Justice; soon afterward he would go out as left vice censor-in-chief and grand coordinator of Shandong.[13] They were neither graduates of the same examination year nor all fellow townsmen, yet they moved in adjoining circles of the Hanlin establishment and the remonstrating offices. In the thirteenth year of Chenghua, Ni Yue, Li Dongyang, and other same-year officials had viewed *Elegant Gathering in the Apricot Garden* in the capital and straightaway commissioned a *Gathering of Hanlin Graduates of the Same Year* in imitation of its precedent. Ni Yue later recalled that this was done "to set in order the affections of our examination year" and to preserve the feeling of meetings and partings, of those living and those gone.[14] No direct causal link can be shown between the bamboo stove exchanges and the same-year gathering picture, yet the two share a single mode of association, in which an object and an image that can be handed round for viewing turn a gathering of a moment into a scroll that can be carried, further inscribed, and remembered.

---

[12] Colophon by Wu Kuan, in *Zhu lu tuyong*, yuan collection, 1a–2a; Li Dongyang 李東陽, "He Sheng Shunchen xinzhi zhu chalu" 和盛舜臣新製竹茶爐, in *Huishan ji* 惠山記, ed. Shao Bao, juan 3.

[13] For the poems and inscriptions of Wu Kuan and those after him, see *Zhu lu tuyong*, yuan collection. On Sheng Yong's career, see Qiu Jun 丘濬, "Duchayuan zuo fu duyushi Sheng gong Yong muzhiming" 都察院左副都御史盛公顒墓誌銘, in *Guochao xianzheng lu* 國朝獻徵錄, juan 60; and Zhang Tingyu 張廷玉 et al., *Ming shi* 明史, juan 162, "Yang Xuan zhuan" 楊瑄傳, with the appended biography of Sheng Yong.

[14] Ni Yue 倪岳, "Hanlin tongnian hui tu ji" 翰林同年會圖記, in *Qingxi mangao* 青溪漫稿, juan 16.

In his colophon to the bamboo stove scroll, Chen Qiong spoke to what such activity meant for the community of scholars. He called the stove "supremely pure and free of the vulgar" and Wang Fu a "pure scholar," and said that Sheng Yu's replicas and Wu Kuan's leading of the exchanges would let "the scholarly grove hear of it with delight and view it with pleasure," and further "transform the customs of scholars, put aside extravagance, and draw near to purity and simplicity, showing that within the teachings of the names there are pleasures of their own."[15] This goes beyond the generic language of refinement. Those who matched poems used the bamboo stove to distinguish purity and simplicity from extravagance, and used one another's inscriptions to confirm which tastes in living accorded with the "teachings of the names." The judgment of an object became at the same time a judgment of persons, and the circulation of the scroll drew the boundaries of a community of scholars pledged to purity and simplicity.

3.1 Sheng Yong's Self-Description as "Bitter Integrity"

If Qin Kui's recovery of the stove aimed chiefly at restoring a piece of local lore, Sheng Yong's naming of the replica stove as "Lord of Bitter Integrity" gave the object a far more definite personal allegory. His inscription reads, "Its form is modeled on Heaven and Earth, though it was neither smelted nor fired. A living fire dwells in its heart; its voice carries the waves of the Xiang. A single drop of sweet dew washes my poet's vitals clean. A clear wind rises beneath my arms, and the eight wilds lie open and bright."[16] The phrase "bitter integrity" comes from the hexagram Jie in the *Book of Changes*, where "bitter restraint; it does not do to persist" originally warned against excessive austerity. In later usage the meaning shifted toward holding fast to one's principles amid adversity, and because bamboo is hollow and jointed and does not change under frost, the phrase attached itself to the nature of bamboo. The bamboo stove is woven of bamboo on the outside and bears the fire within, and when it boils tea it gives off the sound of wind and waves, so that "bitter integrity" came to carry three layers of meaning at once, of form, of use, and of character.

Sheng Yong's official career made this naming something more than a conventional elegant appellation for a tea implement. In 1457, Yang Xuan 楊瑄 impeached Shi Heng 石亨

---

[15] Colophon by Chen Qiong 陳璚, in *Zhu lu tuyong*, yuan collection, 5a–6b.

[16] Sheng Yong, "Kujiejun ming" 苦節君銘, in *Zhu lu tuyong*, yuan collection; see also Gu Yuanqing 顧元慶, *Cha pu* 茶譜, in *Xuxiu siku quanshu* 續修四庫全書, vol. 1115 (Shanghai guji chubanshe, 1995–2002), 218.

and Cao Jixiang 曹吉祥 for encroaching on the people's fields, and Zhang Peng 張鵬, Sheng Yong, and the censors of the thirteen circuits then prepared a joint memorial impeaching the two for their many unlawful acts. When word leaked out, the censors were thrown into prison, and Sheng Yong was demoted to magistrate of Shulu. The *Veritable Records of Yingzong* preserves the course of the imprisonment and the punishments, and the *Ming History* biography of Yang Xuan lists the participants and the exchange at the Wenhua Hall in detail.[17] Qiu Jun 丘濬, writing Sheng Yong's epitaph, went further and claimed that Sheng Yong had first plotted the move with Zhang Peng, holding that the affair touched the fate of the dynasty and asking, "How could we fear calamity and hold our tongues," by which he explained Sheng Yong's "awe-inspiring reputation for straight speech."[18] The full text of the original impeachment no longer survives, and Qiu Jun's account is a retrospective construction, so Sheng Yong's place as the ringleader must be treated with caution. What has firm documentary footing is his demotion for joining the impeachment of powerful ministers, and the use that scholars of the early Hongzhi years made of this episode in shaping his image as a straight-speaking official.

By the time he joined the bamboo stove exchanges in the nineteenth year of Chenghua, Sheng Yong was over sixty and the most senior among the Hanlin juniors. In his own verses he wrote, "When shall I read again the Immortal Lu's lay, where a clear wind of a thousand ages keeps the savor of the Way entire," and again, "This time the idle breeze and moon are roused, and I may divine the joys of return to the fields made whole," closing his final stanza with the lines, "The meat-eaters of this world fall thick as rain; how should they know the fullness of the flavor in my house."[19] The poems carry Lu Tong's 盧仝 song of tea, the Hui spring, and the thought of retiring to the fields, while setting the "meat-eaters" against the pure flavor of his own household. Read together with the name "Lord of Bitter Integrity" and with his early disgrace, the verses show that Sheng Yong was indeed describing himself

[17] *Ming Yingzong shilu* 明英宗實錄, juan 274, entry for the gengzi day of the sixth month in the first year of Tianshun; Zhang Tingyu et al., *Ming shi*, juan 162, "Yang Xuan zhuan."

[18] Qiu Jun, "Duchayuan zuo fu duyushi Sheng gong Yong muzhiming." Qiu Jun writes that Sheng Yong "first plotted with his colleague Zhang Peng of Laishui"; the *Ming Yingzong shilu* records only that "the investigating censors of the thirteen circuits, Zhang Peng and others, joined in a memorial." This article therefore treats Sheng Yong's role as originator as a retrospective account preserved in the epitaph, and not as undisputed fact.

[19] Sheng Yong's three inscribed poems are in *Zhu lu tuyong*, yuan collection, 4b. "The Immortal Lu" in the verses refers to Lu Tong 盧仝.

through the stove. A bamboo stove passes through the fire without being destroyed, and what it lodged was his self-understanding as a man who met setbacks in office without altering his resolve.

This self-description then extended in two directions, through Sheng Yu and through Wu Kuan. The replica stoves were made by the nephew and carried into the capital, and the finished scroll finally came into Sheng Yu's keeping, which gave the whole affair the cast of a family transmission. The participation of Wu Kuan and the other Hanlin and scrutiny officials carried Sheng Yong's self-writing into the association of younger officials. Wang Fu's "lofty integrity," Sheng Yong's "bitter integrity," and Chen Jue's "putting aside extravagance and drawing near to purity and simplicity" answer one another within a single scroll. Together they assert a measure of character acceptable to officials at court, to pass through office without drifting with the vulgar, to meet power with something held in reserve, and to keep the pure pleasures of forests and springs while dwelling within the teachings of the names.

Yet between moral exhortation and a concrete political alliance the evidence still leaves a distance. The surviving titles and colophons name no particular eunuch and no particular policy of the Chenghua court, nor do they state that the company used the bamboo stove to confer on court affairs. Sheng Yong's experience lends "bitter integrity" its political weight; the reverse inference, that every participant in the exchanges shared a hidden political position, cannot be drawn from it. The safer judgment is that this round of exchanges contains a confirmation of values between a senior of the bureaucracy and his juniors of the Hanlin establishment. It could serve as a resource of political attitude, yet it remains something other than a political manifesto that can be reconstructed phrase by phrase.

4. The Bishan Poetry Society and the Space of Mount Hui

The Bishan Poetry Society, founded in 1482, appears at first sight far removed from the official circles of the capital. The ten members, Qin Xu, Lu Mian, Chen Lu, Gao Zhi, Huang Lu, Yang Li, Li Shu, Chen Gongmao, Shi Lian, and Pan Xu, all lived in retirement in the countryside; Pan Xu alone was under sixty, and the rest were mostly elders. The local gazetteers say that in his late years Qin Xu "built a cottage at the foot of Mount Hui" and formed a poetry circle with the venerable elders of the county, and Li Shu's biography lists

his nine fellow members, calling them "all men of high purpose."[20] Zhang Kai's *Record of the Bishan Poetry Society* says that Qin Xu and his companions "respectfully followed the Song men's use of a Tang precedent" in forming the society for their mutual company. So far as the formal membership goes, this was indeed a commoners' society whose conditions of entry were age, poetry, and the life of rural retirement.

Commoner status, however, did not mean the absence of ties to officialdom. Zhang Kai already observed that among these elders "fathers and sons, elder and younger brothers, some were illustrious through the examinations and some renowned for their arts of the Way, so that they were hardly ordinary men of plain cloth."[21] Qin Xu's son Qin Kui served successively as prefect of Wuchang and as administration vice commissioner of Fujian, and was precisely the organizer of the stove's recovery and of the solicitation of poems in the two places. Chen Gongmao's elder brother Chen Mian, a vice director in the Nanjing Ministry of Works, was famed for poetry, painting, and large-character calligraphy. Yang Li came from a wealthy family, and among his brothers were men who rose to vice censor-in-chief and to director of a ministry bureau. Shi Lian and Pan Xu both belonged to Wuxi's hereditary physician families, whose learning was bound up with summonses from the court and service in the Imperial Academy of Medicine. Li Shu himself never held office, yet he was esteemed in his home district for his moral character and literary learning and took part in compiling the Wuxi gazetteer.[22] The members' statuses were far from uniform. What they shared was that, while holding no office, they commanded family learning, literary fame, property, or local standing.

Qin Kui's *Posthumous Manuscripts of the Five Peaks* (Wufeng yigao 五峰遺稿) shows still more clearly that these relations predated the founding of the society. In 1480, on the eve of Qin Kui's journey north, Yang Li invited him together with Huang Lu, Gao Zhi, Chen Gongmao, Shi Lian, Pan Xu, and others for an excursion to the South Lake Thatched

---

[20] *Guangxu Wuxi Jinkui xianzhi* 光緒無錫金匱縣志, juan 25, "Yinyi zhuan" 隱逸傳, entries on Qin Xu 秦旭 and Li Shu 李庶. The sources vary on Qin Xu's courtesy name, writing it Jingyang 景暘 or Jingyang 景昜, and give him the sobriquet Xiujing 修敬; the main text uses his given name only.

[21] Zhang Kai 張愷, "Bishan yinshe ji" 碧山吟社記, in *Huishan gujin kao*, juan 4.

[22] On Chen Mian, see *Kangxi Changzhou fuzhi* 康熙常州府志, juan 30. On Yang Li's family and life, see Wang Ni 王儗, "Changhe shouci houti" 唱和壽詞後題, in *Sixuan wenji* 思軒文集, juan 11. On Shi Lian and Li Shu, see *Guangxu Wuxi Jinkui xianzhi*, juan 25. On the Pan family, see Qin Kui, "Shanggu chushi Pan jun mujieming" 尚古處士潘君墓碣銘, in *Wufeng yigao* 五峰遺稿 (Jiajing-era woodblock ed.), juan 21.

Hall, and Qin Kui addressed the company as "the elders of the poetry altar."[23] The Bishan Poetry Society, in other words, did not gather ten strangers at a stroke. It fixed in place an existing pattern of excursions and verse exchanges. At the recovery of the bamboo stove, Gao Zhi and Lu Mian had already joined the inscribing, and Qin Kui had come and gone repeatedly with Pan Xu, Chen Gongmao, Yang Li, and the rest. When Qin Xu took the head of the society in the eighteenth year of Chenghua, this drew on his years and his moral standing, and it also extended the connections Qin Kui had cultivated into the rural life of the fathers' generation.

This structure also answers the question of why the society admitted fathers, brothers, and men living at home, in place of a society that serving officials might have formed on their own. Officeholders move with their postings and can return home only for mourning, performance evaluation, retirement, or a family visit, which makes a monthly gathering impossible to sustain. The statement in Qin Xu's account of conduct that the Bishan Poetry Society met "once a month" shows that it required a stable local presence.[24] The members at home carried on the daily business of the society, while the officials abroad supplied inscriptions, plaques, introductions, and reputation. The two sides divided the labor within families and within the network of fellow townsmen; neither could replace the other. The society's appearance as a body of "commoners" freed it from the comings and goings of officials, while its influence depended precisely on the members' kin and the Wuxi-born officials who carried its affairs into the wider scholarly world.

The external inscriptions the society received bear this out. Ma Shaorong inscribed the gate with the words "Bishan Poetry Society," Li Yingzhen inscribed the hall as the "Ten Elders," Zuo Zan inscribed the pavilion as "Twisted Whiskers," and Shen Zhou painted the portraits of the ten elders. Shao Bao records that Chen Mian, Sheng Yong, and Qin Kui, though "not among the regular number," took part from time to time, and that the grand coordinator Wang Shu, "hearing of it, came to see them." Li Ying relates further that when Wang Ni and Wu Kuan passed through Wuxi they went about with the elders, until "the canopies of officials' carriages were in sight of one another."[25] What drew the famous men

---

[23] Qin Kui, "You Yang Tingyu Nanhu caotang shi yin" 游楊聽玉南湖草堂詩引, in *Wufeng yigao*, juan 22. The preface records that in the fourth month of the sixteenth year of Chenghua (1480), Yang Li invited Huang Lu, Gao Zhi, Chen Gongmao, Shi Lian, Pan Xu, and others on the excursion.

[24] Qin Kui, "Xianjun Zhenjing xiansheng xingshi" 先君貞靖先生行實, in *Wufeng yigao*, juan 18.

[25] Shao Bao, "Ba Bishan yinshe shijuan" 跋碧山吟社詩卷, in *Huishan gujin kao*, juan 4; Li Ying 李郢,

from outside was a web of relations already under someone's care; no office the ten old men held could have done it. The plaques, poems, and paintings each left a signature behind, and so converted a single visit into a recognition that could be displayed again and again.

The social foundation of the Bishan Poetry Society was therefore neither a covert political group nor merely the self-amusement of retired old men. It took the elders living in the countryside as its daily core, fathers and sons and brothers as its channels between inside and outside, and the Wuxi officials together with the neighboring Suzhou literati as its outer rim of reputation. The members' status as "without office" coexisted with their families' position as "in office," and it was exactly this structure of inside and outside that let the society sustain local life while entering the view of scholars in Jiangnan and in the capital itself.

4.1 Xiangshan and White Lotus Precedents and the Bounds of Archaistic Meaning

The Bishan Poetry Society's account of itself cannot be separated from ancient stories of fellowship. In his colophon to the *Painting of the Bishan Poetry Society*, Shen Zhou first cited the White Lotus Society of Huiyuan, writing, "The Reverend Yuan of Mount Lu founded the Lotus Society, and Yuanming and Xuandu were in its company; this was the glory of a thousand ages," and continued, "The ten elders of Liangxi have imitated them and built a hall on the sunny slope of Mount Hui."[26] The precedent of the White Lotus Society links monastery, mountains and forests, scholars, and monks, and it suits the setting of the Listening to the Pines Hermitage perfectly. Wang Da had already used the phrase "the Reverend Yuan planting lotuses" in his *Preface to Pure Poems on the Bamboo Stove*, and Wang Fu's poem on the stove has the line "the old Chan monk entrusted it with reopening the society." The White Lotus Society was thus more than a generic allusion to lofty men. It could join the early Ming exchanges of Xinghai and Wang Fu at the Listening to the Pines Hermitage with the Bishan Poetry Society of the Chenghua years.

The prefaces and colophons of the society nonetheless draw more often on the societies of Xiangshan and Luoyang. Zhang Kai placed Bai Juyi's 白居易 Nine Elders of

---

"Bishan yinshe yaji tu ji" 碧山吟社雅集圖記, in the same juan.

[26] Shen Zhou 沈周, colophon to *Bishan yinshe tu* 碧山吟社圖, transcribed in *Huishan gujin kao*, juan 4. The transcription reads, "The Reverend Yuan of Mount Lu founded the Lotus Society, and Yuanming and Xuandu were in its company . . . The ten elders of Liangxi imitated them and built a hall on the sunny slope of Mount Hui, naming it the Bishan Poetry Society."

Xiangshan alongside the ten elders of Bishan, the former living "at the moment of merit accomplished and self withdrawn," the latter "singing of peace in days when those above are tranquil and those below serene." Qin Kui likewise wrote in a poem presented to the company, "Their talents suffice to carry on the glory of Xiangshan," adding, "I too am weary now of wandering in office, and would gladly keep company with the yellow-haired among fishermen and woodcutters."[27] The emphasis here falls first on great age, second on withdrawal from office, and third on the pleasures of poetry and wine. The Xiangshan precedent gave an orderly name to a local gathering whose members differed in age and station, and it also accommodated the imagination of serving officials about life after return to the fields.

The *Painting of the Bishan Poetry Society* by Shen Zhou lifted the ten elders out of the scene of their monthly meetings, preserving their features, their seats, and the surroundings of Mount Hui, while his colophon placed this local gathering within the succession of fellowships through the ages. Wang Yu, examining Shen Zhou's associations at Mount Hui, has noted his dealings with Qin Kui, which provides a traceable line for the genesis of the painting.[28] Shen Zhou's involvement in the Wuxi society's affairs thus came through Qin Kui and through the local network of painting and calligraphy; it was no intervention out of nowhere. Since Zhang Kai already identified the story of the Nine Elders of Xiangshan as the model the Bishan Poetry Society imitated, the historical circumstances of that model demand our attention. In the fifth year of the Huichang era of Emperor Wuzong of Tang (845), Bai Juyi resigned his office as Minister of Justice, bought the grounds of the former Xiangshan Monastery within the city of Luoyang, and gathered there with eight friends in elegant assembly, forming what was called the Society of the Nine Elders, whence the name "Nine Elders of Xiangshan." His reason for resigning and retiring was to escape a perilous court in which eunuchs wielded power and factional strife never ceased.

The *Painting of the Bishan Poetry Society* was very likely made under the inspiration of *Elegant Gathering in the Apricot Garden*. In 1477, after viewing the Apricot Garden scroll in the capital, Ni Yue and his companions asked a painter to make the *Gathering of Hanlin*

---

[27] Zhang Kai, "Bishan yinshe ji"; Qin Kui, "He Bishan yinshe xincheng yun" 和碧山吟社新成韻 and "Zeng Bishan yinshe zhugong" 贈碧山吟社諸公, in *Wufeng yigao*, juan 8.

[28] Wang Yu 王瑀, *Shen Zhou shihua zhong de Jiangnan* 沈周詩畫中的江南 [Jiangnan in Shen Zhou's poetry and painting] (Beijing daxue chubanshe, 2022), 138, 149.

Graduates of the Same Year, as Ni Yue's own record attests.[29] Wu Kuan, Li Dongyang, and others afterward organized repeated gatherings of fellow townsmen and of men of the same examination year. This shows that scholars of the Chenghua period were familiar with the commemorative sequence of gathering, portraits, inscriptions, and the making of a scroll. Shen Zhou never saw the Apricot Garden painting himself, yet he could well have learned of it, of the Hanlin same-year picture, and of the deeper meanings behind them through his close friend Wu Kuan. Because Shen Zhou had seen Song versions of *The White Lotus Society* and *The Nine Elders of Xiangshan* in the collecting circles of Jiangnan, his painting drew its formal vocabulary, in all likelihood, from those Song works, and any debt to *Elegant Gathering in the Apricot Garden* remains unproven. Whether or not Shen Zhou knew of the political expression hidden in the Apricot Garden and the same-year pictures and put it to use in his own painting, for the local gentry of Wuxi such pictures had already become one available way of voicing their own ideals.

4.2 Objects, Scrolls, and Paintings in the Shared Fashioning of Mount Hui

Read together, the bamboo stove and the poetry society show that the space of Mount Hui was no static backdrop to cultural activity. It was affirmed again and again in the return of the object, the deposit of the scrolls, the building of the society's quarters, and the making of the painting. The Listening to the Pines Hermitage kept both the old poems and the new poems of the recovered stove. The Second Spring supplied the irreplaceable water for brewing tea. The Bishan Poetry Society set its gate, hall, and pavilion beside the spring. Object and buildings thus shared a single stock of place names, persons, and stories. Even where poems were produced in Beijing, the finished scrolls pointed back to Mount Hui. Even though Shen Zhou seldom kept company with the ten elders, what his picture painted took the society's gatherings at Mount Hui as its subject.

On obtaining the stove, Qin Kui returned it to the Listening to the Pines Hermitage, declining to keep it in his own household, and when the capital scroll was complete it too was given to the monks for safekeeping. The return of the old object and of the poem scroll turned a mobilization of personal obligation into a local monument. Sheng Yu's replicas show further that this meaning no longer belonged to one original object alone. So long as the form,

[29] Ni Yue, "Hanlin tongnian hui tu ji." The record states explicitly that in the dingyou year of Chenghua someone in the company produced the *Elegant Gathering in the Apricot Garden* scroll, whereupon those present had a painter make their portraits into a painting.

the naming, and the web of poems endured, the bamboo stove could enter new spaces and new communities. Once the poems and portraits had been arranged, mounted, and inscribed afresh, they lasted far longer than any single tea gathering or society meeting, so that Shao Bao, Qian Fu, and Lu Jian in later years could still append prefaces and colophons and reinterpret what had gone before. These objects and scrolls moved from hand to hand, yet always with the Listening to the Pines Hermitage, the Second Spring, and the Bishan quarters as their destination, and Mount Hui grew from a celebrated spring and scenic spot into a literati space possessed of a genealogy.

In this process, local reputation and personal reputation drew strength from one another. Qin Kui used Wang Fu's name to recover the stove, and in recovering it and soliciting poems he left behind a local accomplishment of his own. Sheng Yong used the "Lord of Bitter Integrity" to declare his aspirations, while the stove gained its meaning of steadfastness from his official career. The ten elders entered the wider scholarly world through the plaques, poems, and paintings of famous outsiders, and those who made the presentations used the old precedents of Xiangshan and the White Lotus to lodge their own imaginings of withdrawal. Zhang Kai's dictum that "a thing does not make itself precious" and Chen Qiong's picture of "the scholarly grove hearing with delight" together reveal that reputation belongs to no object or person singly. It is generated in repeated inscription and mutual judgment.

This web of relations also shows why the phrase "local cultural capital," though it describes the outcome well enough, cannot explain the process. The participants were doing more than adding famous stories to Wuxi; they were managing their own situations. Men in mourning needed to settle a life temporarily withdrawn from office within their home district. The retired needed to preserve the renown of their years in office. Those still serving used the imagination of return to the fields to balance court and marketplace against forests and springs. Commoner physicians and scholars used the society's gatherings to confirm their standing in the locality. Mount Hui could hold these needs together exactly because it lay near the city and its routes of travel while also possessing a famous spring, monasteries, and the traces of former worthies, a place where local life and the experience of the court could meet in one spot.

Seen in this light, the political meaning of these undertakings does not lie in hunting out a common program concealed behind the lines of verse. What the bamboo stove and the poetry society maintained was an order of judgment about who could be a friend, what manner of life could be called pure, and how a man should conduct himself between office

and withdrawal. Such judgments were closely tied to experience of officialdom, and at particular moments they could become grounds for criticizing the powerful. The evidence in Sheng Yong's case is relatively strong. Qin Kui's recovery of the stove and the founding of the Bishan society chiefly display the organization of a local cultural lineage and of a network of association. Only by distinguishing these several levels can we see both that cultural activity was never unconnected to politics and that refined amusement cannot all be written as cipher.

5. Conclusion

The exchanges on the recovery of the bamboo tea stove and the founding of the Bishan Poetry Society in Chenghua-era Wuxi appear on the surface as two local cultural events, one of a stove and one of a society, yet they shared a single network of persons and a single stock of symbolic resources. The narrative of the bamboo tea stove was rewritten continually over several decades. Wang Fu, one participant among the original company of inscribers, was pushed step by step into the position of the stove's maker, and his persona of "lofty integrity beyond the vulgar," built on the steely and unbending groundwork of the biographies written by his contemporaries and exalted through Ye Sheng's anecdote of the flute heard beneath the moon, was bound fast to the stove as an object, so that it became a spiritual token that locality and family alike could hand down. Through the recovery and the exchanges, Qin Kui restored a local monument and prolonged the cultural line of Mount Hui. Sheng Yong, through the inscription "Lord of Bitter Integrity," lodged his resolve in the object, taking his own experience of impeaching Shi Heng and Cao Jixiang and enduring demotion without yielding as his ground, and conducted with Wu Kuan, Li Dongyang, and the other Hanlin lecturers who would soon hold the future of the court in their hands a transmission of spirit across the generations. The ten commoner elders of the Bishan Poetry Society, through a network in which their sons, brothers, and marriage kin mostly held eminent posts, drew an unbroken stream of presentations from the leading figures of an age, Wang Shu, Wang Ao, Wu Kuan, and the rest, so that the classical ideal of Xiangshan and Luoyang was practiced by a circuitous path, in the form of a local society, beneath the realities of eunuch usurpation.

Returning here to Yin Ji'nan's observation about *Elegant Gathering in the Apricot Garden*, that the gathering pictures of Beijing officials in the mid-fifteenth century were in fact "a political painting deeply harboring the antagonistic complex between civil officials

and eunuchs," the Wuxi case shows that this undercurrent existed beyond the gatherings of capital officials and reached into local society. The bamboo stove, the poetry society, and the aspiration toward forests and springs were at once the cultivation of cultural capital and, in a context of disordered court government, the means by which a gentry community sustained its shared values, admonished its juniors, and curated local memory. The Second Spring of Mount Hui won empire-wide fame through Lu Yu's grading of waters, and through the bamboo tea stove and the Bishan Poetry Society it became a cultural space written and overwritten, layer upon layer, by the gentry network. The three scrolls of the *New Poems on the Bamboo Stove* are kept at the Listening to the Pines Hermitage and with the Sheng family, while a single painting of the Bishan Poetry Society has circulated in the world of art. Object, text, image, and place here made one another whole.

Bibliography

Guangxu Wuxi Jinkui xianzhi 光緒無錫金匱縣志.

Gu Yuanqing 顧元慶. *Cha pu* 茶譜. In *Xuxiu siku quanshu* 續修四庫全書, vol. 1115. Shanghai guji chubanshe, 1995–2002.

Hu Guang 胡廣. *Hu Wenmu gong wenji* 胡文穆公文集.

Kangxi Changzhou fuzhi 康熙常州府志.

Li Ying 李郢. "Bishan yinshe yaji tu ji" 碧山吟社雅集圖記. In *Huishan gujin kao*, juan 4.

Ming Yingzong shilu 明英宗實錄.

Ni Yue 倪岳. *Qingxi mangao* 青溪漫稿.

Qin Kui 秦夔. *Wufeng yigao* 五峰遺稿. Jiajing-era woodblock edition.

Qiu Jun 丘濬. "Duchayuan zuo fu duyushi Sheng gong Yong muzhiming" 都察院左副都御史盛公顒墓誌銘. In *Guochao xianzheng lu* 國朝獻徵錄, juan 60.

Shao Bao 邵寶. "Ba Bishan yinshe shijuan" 跋碧山吟社詩卷. In *Huishan gujin kao*, juan 4.

Shao Bao 邵寶. *Rongchuntang ji* 容春堂集.

Shao Bao 邵寶, ed. *Huishan ji* 惠山記.

Shen Zhou 沈周. Colophon to *Bishan yinshe tu* 碧山吟社圖. In *Huishan gujin kao*, juan 4.

Tan Xiu 談修, comp. *Huishan gujin kao* 惠山古今考. Wanli-era woodblock edition.

Wang Da 王達. "Zhu lu ji" 竹爐記. "Zhu lu qingyong xu" 竹爐清詠序. In Wu Yue, *Zhu lu tuyong*.

Wang Hong 王洪, and Zhang Bingru 章昞如. Appendixes to *Wang sheren shiji* 王舍人詩集. In *Xishan xianzhe congkan* 錫山先哲叢刊, vol. 2. Zhonghua shuju, 1922.

Wang Ni 王儗. *Sixuan wenji* 思軒文集.

Wang Yu 王瑀. *Shen Zhou shihua zhong de Jiangnan* 沈周詩畫中的江南. Beijing daxue chubanshe, 2022.

Wu Yue 吳鉞, comp. *Zhu lu tuyong* 竹爐圖詠. Vol. 4 of *Xishan xianzhe congkan* 錫山先哲叢刊. Zhonghua shuju, 1922.

Ye Sheng 葉盛. *Shuidong riji* 水東日記. Zhonghua shuju, 1980.

Yin Ji'nan 尹吉男. "Zhengzhi haishi yule: Xingyuan yaji he Xingyuan yaji tu xin jie" 政治还是娱乐：杏园雅集和《杏园雅集图》新解. *Gugong bowuyuan yuankan* 故宫博

物院院刊, no. 1 (2016): 6–39.

Zhang Kai 張愷. “Bishan yinshe ji” 碧山吟社記. In *Huishan gujin kao*, juan 4.

Zhang Tingyu 張廷玉 et al. *Ming shi* 明史.

Research Paper

Feminism in Translation: Sociopolitical Context, Feminist Cinema, and Chinese Reception of *Barbie, Portrait of a Lady on Fire, and Night and Fog*

Ziyue Li
Grade 12

Editors' Introduction

This paper presents a compelling exploration of how sociopolitical contexts shape feminist cinema and how feminist ideas are reinterpreted by Chinese audiences. Its greatest strength lies in its interdisciplinary approach, integrating feminist theory, film studies, and cross-cultural reception analysis.

Nevertheless, several aspects could be strengthened. The historical overview occasionally relies on broad distinctions between American, French, and Hong Kong feminism, overlooking complexities within each tradition. Additionally, although the reception analysis incorporates valuable primary materials, the selection of online comments and reviews lacks a sufficiently systematic methodology, making some broader conclusions difficult to substantiate. Certain sections also prioritize description over critical interpretation, and occasional repetition affects the overall precision of the prose.

Overall, this is an ambitious and intellectually engaging paper that demonstrates strong interdisciplinary awareness and independent thinking. Future research would benefit from a more rigorous approach to audience reception, closer formal analysis of cinematic techniques, and greater attention to alternative interpretations.

1. Introduction

The aim of this paper is to explore how sociopolitical context in the United States, France, and Hong Kong has shaped feminist filmmaking, and how these varied feminist ideas are received and reinterpreted by Chinese audiences.

This essay examines the intersection of film, feminism, and cultural context through three films: Greta Gerwig's *Barbie* (US, 2023), Céline Sciamma's *Portrait of a Lady on Fire* (France, 2019), and Ann Hui's *Night and Fog* (Hong Kong, 2009). Gerwig, an American writer and director famous for female-centered narratives, uses humor together with the commercial language of Hollywood to satirize patriarchy and the long-standing expectations imposed on women. Sciamma, a French director whose work focuses on female identity, queer identity, and desire, constructs an ideal world away from "male gaze", delicately depicting how two women look, help, and love each other. Ann Hui—a key figure of the Hong Kong New Wave—explores how women express themselves – or endure in silence – in ordinary scenes of daily life, together with larger social issues like immigration and family structures. Although these three films differ in theme, genre, and cinematic form, the differences offer a useful entry point for exploring how feminism manifests distinct cinematic forms in different cultural contexts.

Despite my own Chinese nationality, mainland China is not considered as a fourth context in this paper, because it has a more strictly regulated film industry. Films must pass state review for public release and works that relate to gender, sexuality, or other socially sensitive issues may have to negotiate and adjust many times to meet exhibition limits. Although more female filmmakers have entered the Chinese film industry in recent decades, women remain underrepresented in many major creative and decision-making positions. Therefore, this paper treats mainland China primarily as a context of reception. By examining how *Barbie*, *Portrait of a Lady on Fire*, and *Night and Fog* are viewed and commented upon by Chinese audiences, I hope to understand how feminist ideas travel across cultural boundaries: which ideas resonate, and how they change when they enter a different social and cultural environment. This question is also relevant to my own position as an aspiring filmmaker in China. I consider it in light of the China Film Administration's film policy, Shaw and Zhang's scholarship, and the University of Amsterdam's discussion of Chinese female filmmakers (Sun and Xu).

The first section gives an overview of the development of feminism in the United States, France, and Hong Kong, and considers how these different cultural histories have

influenced feminist filmmaking. The second section analyzes *Barbie*, *Portrait of a Lady on Fire*, and *Night and Fog*, focusing on the different cinematic languages through which they express feminist concerns. The final section turns to their reception in China.

Because the three films entered Chinese culture in very different ways, I use different kinds of sources to study each one. *Barbie* arrived as a mainstream commercial film and generated large online debates, so I examine highly engaged Bilibili videos and comments, together with media reports and academic studies. *Portrait of a Lady on Fire* circulated more through film festivals, cinephile, feminist, and queer film spaces, so I use material from the Beijing International Film Festival, long-form Douban and Bilibili reviews, and discussions by Chinese critics and scholars. For *Night and Fog*, a Chinese-language Hong Kong film, I examine mainland reviews and online videos together with discussions of Hong Kong society and interviews with Ann Hui.

2. Three Feminisms

Although women in all three contexts have challenged gender inequality, the questions they asked and the actions they took were distinct due to their own political systems, social structures, cultural traditions, and timeline of feminist developments. In other words, feminism did not develop in the same way everywhere, and these different cultural values and political moments made different problems more urgent for women in each society, and also shaped the ways they responded to them.

In the United States, feminism is often described as having developed in four "waves." The first wave, beginning in the nineteenth century, focused primarily on women's legal status and basic political rights, including property, education, and suffrage. Early American feminists asked a simple but fundamental question: could women participate in society as full legal and political citizens? The ratification of the Nineteenth Amendment in 1920 was an important achievement of this period. By the 1960s and 1970s, however, many feminists had begun to argue that equal rights on paper did not necessarily mean equality in everyday life. The slogan "the personal is political" reflected a shift toward relating the structures of patriarchy with marriage, family, education, work, sexuality, and the media. Issues like reproductive freedom, unpaid domestic labor, and cultural representation became central. Feminist film theory developed within this context. Beyond asking whether women were represented positively, critics began to ask about conventional cinematic language itself: was it structured around a male spectator? Take Laura Mulvey's concept of the "male gaze"

for example; she thought that mainstream cinema often positions men as active subjects of action and vision, while women become images whose primary function is to be seen (Alexander; National Archives; Munro; Mulvey 11).

Later developments in American feminism further contributed to the global conversation, moving away from a generic idea of "women" toward a more nuanced understanding of women's different experiences. Third-wave feminism in the 1990s and 2000s emphasized diversity, identity, sexuality, and intersectionality. Instead of asking only how "women" as a single group were represented, feminists increasingly questioned inside the definition of the "group": which women had historically been included, and whom excluded, in that category? How does gender interact with race, class, sexuality, nationality, and other forms of identity? Several influential writers helped develop these questions. Audre Lorde and Bell Hooks refused to assume all women experienced oppression in the same way, holding the opinion that feminism should take racial, class, and sexual differences among women seriously. Kimberlé Crenshaw's concept of "intersectionality" later provided a framework for understanding how different forms of inequality can overlap, while Judith Butler's *Gender Trouble* (1990) questioned the idea that gender itself is a fixed and natural identity (Lorde; hooks; Crenshaw; Butler).

Fourth-wave feminism carries many of these debates into digital culture. Online activism, #MeToo, discussions of consent, body autonomy, sexual violence, and the representation of marginalized identities became easier to circulate publicly. Meanwhile, the film industry itself has been criticized: even though women appear on screen in a greater diversity of roles and narratives, who is allowed to direct, write, produce, and control the images? Whose story will be told? These challenges are not without reason. *Half the Picture* (2018), a documentary about gender inequality among film directors, points out that women and men attend film schools in roughly equal numbers, yet only about 4% of the highest-grossing films in the United States over the previous decade were directed by women. Changes in the industry have been modest in the years since: in 2025, just 8.1% of directors of the top 100 grossing US theatrical films were women. (Munro; *Half the Picture*; Smith et al.; Gonzalez).

At the same time, some recent films greenlit for women directors suggest a change in content: women's perspectives and feminist concerns are becoming more explicit. Take some American films I've watched for example: *Little Women* (2019) and *Barbie* (2023), both directed by Greta Gerwig, place women at the center of the story and also of its authorship. Similarly, *Promising Young Woman* (2020), directed by Emerald Fennell, tells a story about

sexual violence and consent from a woman's perspective rather than building the narrative around the male perpetrator. And Maggie Gyllenhall's *The Bride* (2026) takes "one of the most famous novels by a 19th-century female writer" (Mary Shelley's Frankenstein) to "reconsider the vexed, enduringly provocative figure of the monstrous woman" (Dargis). It reworks a familiar story, that of "Frankenstein's bride" (an idea which Shelley never explored but which emerged in cinematic spinoffs from the novel) from a woman's perspective, placing the Bride herself at the center rather than treating her as an extension of the male monster. These films show that fourth-wave feminist cinema is concerned not only with whether women appear on screen, but also whose experiences are taken seriously and who has the power to frame the story.

In France, by contrast, the development of feminism is harder to divide into clear stages. One important starting point was Simone de Beauvoir's *The Second Sex* (1949), which questioned the idea that femininity was something natural and instead examined how women were socially shaped by society as the "other". Later, especially around the feminist movements of the late 1960s and 1970s, such as the Mouvement de Libération des Femmes (MLF) around 1970 and the the "Manifesto of the 343" in 1971, French feminists pushed for reproductive rights, sexual freedom, and greater control over women's own bodies (Beauvoir; Library of Congress, "20th Century").

Feminists also asked how femininity was produced through language, culture, and representation rather than treating gender simply as a legal category. This manifests an important difference from the American trajectory. American feminist debates often moved very visibly through questions of rights, structural patriarchy, intersectionality, and public activism. In France, questions about women's freedom were also closely connected to how women understood their bodies, desires, identities, and positions as subjects. Beginning in the 1970s, French female philosophers, literary theorists, and psychoanalysts played an important role in developing these ideas. Writer and literary theorist Hélène Cixous developed the idea of écriture féminine, or "feminine writing", which explores whether women could express their bodies and experiences through writing not organized around male experience; philosopher and psychoanalytic thinker Luce Irigaray questioned male-centered philosophy and psychoanalysis on the grounds that it left little space for female subjectivity; and literary theorist and psychoanalyst Julia Kristeva examined language, desire, the unconscious, and the maternal body. They all questioned how women could have become subjects of speech and desire within a cultural system linguistically and symbolically constructed around men. French feminists were interested in whether women could become subjects who look and

desire, rather than simply objects to be looked at (Cixous; Irigaray; Kristeva).

These concerns continue in contemporary French feminist filmmaking. In the French films I studied over the course of my research, including Agnès Varda's *Cléo from 5 to 7* (1962), Chantal Akerman's *News from Home* (1976), Noémie Merlant's *The Balconettes* (2024), and Céline Sciamma's *Water Lilies* (2007), *Petite Maman* (2021), and *Portrait of a Lady on Fire* (2019), female subjectivity is often centered in how women feel, how they view the world, and how they express desire. This is especially clear in Céline Sciamma's work. Her films often focus on female identity, desire, intimacy, silence, and observation, and emotions are frequently expressed through eyes, bodies, space, and small gestures rather than direct explanation. Other female filmmakers explore similar questions in different ways. Noémie Merlant explores female anger, violence, trauma, and patriarchy through women's bodies, emotions, and performances, making women's own experiences the center of the film; Julia Ducournau uses the body itself to question social expectations; and Justine Triet presents women as complicated and sometimes ambiguous instead of making them victims or heroes, with complex desires, perspectives, and contradictions.. In different ways, these filmmakers carry forward the French feminist concern with whether women can appear as subjects who feel, look, desire, and speak for themselves, rather than simply as objects defined by others.

In comparison with the more public-facing and direct feminist language in the United States, French feminist filmmaking often places more attention on the body, desire, subjectivity, and the act of looking. The connection between French feminist history and its feminist cinema, then, is that feminist questions are included in the form of the film itself: whose desire organizes the story, who is allowed to look, and from whose perspective a woman's experience is shown. This background is particularly important for understanding *Portrait of a Lady on Fire*, which will be discussed in the following section.

Hong Kong feminism developed in a very different social environment. The rise of Hong Kong feminism intersects in complex ways with Hong Kong's evolving economic, cultural, and political pressures. During the rapid industrialization of the 1970s and 1980s, more women entered the labor market. This brought greater attention to problems such as wage inequality, occupational discrimination, labor rights, and legal equality. By the 1980s and 1990s, the discussion had expanded. Domestic violence, sexual assault, and other structural forms of gender inequality became more prominent concerns. Women's organizations and the development of civil society helped bring these problems into public discussion (Lim).

The 1997 handover, an important political context of Hong Kong that made its feminist development different from that of the United States and France, added another layer to the feminist questions above. Gender became connected with identity, migration, belonging, and Hong Kong's changing relationship with both Britain and mainland China. Eliza Lee describes women's experiences in Hong Kong as being shaped through the interaction of globalization, postcolonial conditions, and Chinese patriarchy. This combination created internal conflict and tension . More women received higher education and gained professional jobs, but expectations inside the family did not change at the same speed. Women in the 90s and early 2000'scould work full-time and still be expected to take care of the household, children, and other family responsibilities. In other words, gaining more freedom in public life did not necessarily remove inequality in private life. My earlier research describes this as a "double burden": professional work on one side and unpaid care work and domestic labor on the other (Lee).

After 1997, Hong Kong feminist discussions, influenced in part by the circulation of works in translations by writers such as Judith Butler, Julia Kristeva, and Gayatri Chakravorty Spivak, increasingly included LGBTQ identities, non-traditional families, migration, and varied forms of female desire and identity. Instead of simply importing Western feminism, Hong Kong scholars and activists had to consider how these ideas changed when they encountered Hong Kong's own questions of colonial history, Chinese culture, migration, family, and sexuality. Adelyn Lim's research on post-1997 Hong Kong emphasizes that feminist activism involved groups of women with very different social and cultural positions. This made it necessary to speak about "Hong Kong women" not as one uniform group but as a diverse group in which gender intersects with class, migration status, sexuality, and cultural identity, while family remains an important space in which these tensions were negotiated (Lim).

These concerns are reflected in Hong Kong feminist filmmaking, which often focuses on ordinary life: marriage, family, work, care, migration, and survival, rather than only in explicit political confrontation, like that in the United States and France. My research on Hong Kong filmmaking finds a recurring theme: how women seek autonomy while remaining restricted in those ordinary relationships and obligations. Ann Hui is particularly relevant to this tradition. Her films, such as *The Way We Are* (2008), *Night and Fog* (2009), and *All About Love* (2010), often focus on women and other marginalized groups, and female anger may appear through silence, endurance, or survival rather than direct confrontation. In *The Way We Are*, a widowed mother works in a supermarket, takes care of her teenage son,

and gradually builds a friendship with an elderly neighbor. Her strength is shown through everyday work, care, and endurance. In *All About Love*, two former female lovers reunite while both are pregnant, and the film uses their relationships, motherhood, and daily lives to question traditional ideas about family and sexuality. This does not mean that endurance is feminist; instead, it shows how Hong Kong's particular history and social structure have shaped a different way of thinking about women's freedom, a context in which women's choices are closely tied to family, economic dependency, migration, and social institutions. These conditions provide an important background for *Night and Fog*, the Hong Kong feminist film that I'm going to discuss (*The Way We Are*; *Night and Fog*; *All About Love*).

3. Three Films

The three films discussed in this paper express feminism in very different ways. Greta Gerwig's *Barbie* uses comedy, fantasy, and direct political language; Céline Sciama's *Portrait of a Lady on Fire* builds its feminist perspective through looking, desire, and intimacy; Ann Hui's *Night and Fog* places women's struggles within family, migration, poverty, and domestic violence. Their differences are not only in what they talk about, but also how their stories are told, how the camera looks at women, and which emotions women are allowed to express. In this sense, the three films reflect the different forms of feminism in the United States, France, and Hong Kong.

In *Barbie*, feminism is expressed through its story directly. *Barbie* is a doll who lives in Barbieland, a fantasy world inhabited by multiple versions of *Barbie* and Ken. At the beginning of the film, Barbieland appears as a perfect female world: contemporary Barbies can be presidents, doctors, writers, and judges, while the Kens, male, can only exist around them. However, when *Barbie* enters the real world, she realizes that this equality is only imagined; women's actual experiences are not that ideal. Ken, meanwhile, discovers patriarchy in the real world and brings it back to Barbieland. The social order in Barbieland is quickly reversed. Accordingly, *Barbie* turns a complicated gender system into a very visible and exaggerated conflict. There, patriarchy is clearly named rather than hidden. This directness reflects the form of American feminism discussed earlier, which has taken a highly public form through popular culture, social media, and movements such as #MeToo. Since some audiences may not understand feminist theory wholly, *Barbie* turns ideas about patriarchy and gender expectations into jokes, speeches, songs, and bright visual symbols. What is distinctive about *Barbie*, then, is not just its "feminism", but that it turns feminism

into the language of mainstream popular culture, a Hollywood blockbuster (*Barbie*).

Barbie makes its concerns very specific. Toward the close of the film, Gloria, a mother who, together with her 13-year-old daughter Sasha, pursues *Barbie* in an extended chase scene through the "real world" to rescue her from the machinations of the Mattel company, gives a speech about the contradictory expectations placed on women: they are expected to be attractive but not threatening, confident but not arrogant, and independent while still caring for others around them. After the Kens take control of Barbieland, many of the Barbies begin to accept roles centered on pleasing and supporting men. Gloria's speech helps them recognize that these expectations which limit their choices have started to feel normal. Like Gloria, the Barbies come to understand that they have internalized a patriarchal structure, and they decide to work together to regain control of Barbieland (*Barbie*).

Ken's story shows the problem from the men's side. After discovering patriarchy in the real world, he tries to take this hierarchy that gives men more rights than women back to Barbieland. He calls on other Kens to build their identity around status, dominance, possessions, and control over the Barbies. His exaggerated transformation demonstrates how toxic masculinity can come from the pressure on men to prove themselves through power. By the end, Ken also has to imagine who he is without controlling or being defined by *Barbie*. By making its feminist concerns explicit, *Barbie* made these issues visible to a much wider audience. Its direct and exaggerated style created a lot of controversy, but this controversy also brought attention, an indispensable foundation for any progress. I'll discuss the controversy centering *Barbie* in Part 4 (*Barbie*).

Beyond the story, the film's feminist perspective can be found in the narrative and growth of its principle character, *Barbie*, as she gradually becoming aware of herself as more than an image. *Barbie* begins as a doll whose identity depends on "looking perfect". Her body, clothes, house, and daily routine are all designed to be seen. But as the story develops, she begins to experience interiority: discomfort, aging, and therefore fear and sadness. These changes are natural as people become older, but *Barbie* cannot adjust to this uncertainty. Through these unfamiliar experiences, *Barbie* gradually learns that her value does not depend on remaining perfect and that uncertainty is part of being human. She can accept uncertainty and decide for herself what kind of life she wants. This is also why the ending matters: instead of simply restoring the original female-dominated Barbieland or returning to her old life, she chooses to become human. Therefore, the feminist value of *Barbie* goes beyond the idea that women should "win" the "battle of the sexes" that the film so overtly stages: it reveals that women should be allowed to exist beyond fixed roles and impossible

expectations. The film also demonstrates this idea through its focus on motherhood. The Mattel *Barbie* doll originally offered girls an alternative to baby dolls and the expectation that they should imagine themselves mainly as mothers, but the film reclaims motherhood: Gloria and Sasha's relationship is shown not only through traditional maternal care; they also act as companions and partners throughout the story, especially as they follow *Barbie* through the real world and later help her reclaim Barbieland. At the same time, *Barbie*'s final choice to become human brings her into a world of real bodies, aging, and the possibility of reproduction. In this sense, the film both reclaims and redefines motherhood (*Barbie*; Library of Congress, "Ruth Handler").

Portrait of a Lady on Fire expressed its feminism very differently. The film does not point out patriarchy directly and does not use many speeches to explain women's oppression, like in *Barbie*. Instead, its feminist perspective is built into the relationship between Marianne, a female painter, and Héloïse, the woman she is asked to paint. Marianne has been commissioned to paint a portrait of Héloïse to be sent to a man she has never met, so he can decide whether to marry her. At first, Marianne must secretly observe Héloïse because Héloïse refuses to sit for a marriage portrait. Héloïse is told that Marianne has been hired as her walking companion; Marianne studies and memorizes Héloïse's face and movements during their walks and paints her from memory at night. In this situation, looking is a form of power: Héloïse is being turned into an image for a male viewer (*Portrait of a Lady on Fire*).

After Héloïse discovers Marianne's real purpose, however, their relationship begins to change. She agrees to sit for the portrait, but she also looks back. She criticizes Marianne's first portrait and points out that while Marianne has been studying her expressions and movements, she has also been studying Marianne. Héloïse is still a "model", but she is not a passive object, because the painter no longer controls the gaze; it is shared. Later, the second portrait becomes more collaborative because both women participate in the process of looking and representation (*Portrait of a Lady on Fire*).

This mutual looking is one of the clearest ways the film creates a female gaze. Men are almost absent from the film, although male power still exists through Héloïse's arranged marriage and the restrictions placed on Marianne as a woman artist. Sciamma creates a temporary space in which women can look at one another, desire one another, make art, and care for one another without men's control. She has described the film as a love story built around equality and has rejected the traditional division between the artist and the "muse"; for her, the model and the artist are co-creators. In diametric opposition to the direct feminist language of *Barbie*, *Portrait of a Lady on Fire* expresses its politics through silence, eye

contact, bodily distance, painting, and memory. Its tone is restrained and intimate rather than openly comic or angry (*Portrait of a Lady on Fire*; Stevens).

Moreover, even when the two women cannot remain together in the end due to realistic reasons, the film does not depict their relationship as a tragedy. Their love continues through memory, art, and the experience of having seen and been seen by one another. Critics have also emphasized that the film allows its feminist perspective to emerge through the way women see and represent one another. In this way, *Portrait of a Lady on Fire* reflects the French feminist concerns discussed in the previous section: female subjectivity, desire, the body, and the question of who has the right to look (*Portrait of a Lady on Fire*).

Night and Fog takes yet another approach. Inspired by a real domestic-violence case, the film follows Wong Xiaoling, a mainland Chinese woman who moves to Hong Kong after marrying Lee Sum, an older Hong Kong man. They live with their two young daughters in Tin Shui Wai, where their marriage unravels, as Xiaoling's husband gradually becomes violent and controlling. Xiaoling tries more than once to leave and seek help, but she is trapped inside the structures that shape her choices. Many realistic factors make escape difficult: her immigrant status, motherhood, joblessness, lack of social support. Again and again, she has to return to her husband, even though he often beats her hard, hurting both her and their children. In the end, she and her children are killed by him. The film tells their story through flashbacks, slowly revealing how the relationship develops and how Xiaoling becomes trapped in it (*Night and Fog*).

The real case behind the film makes this even clearer: the woman had repeatedly sought help, had registered for employment training as she prepared for a more independent life, and had gone to a police station on her final day. She was trying to change her situation, but her choices were limited by the conditions around her (*Night and Fog*; "许鞍华｜谈电影").

Xiaoling's situation makes the feminist question in *Night and Fog* very different from that of the other two films. Ann Hui does not create an imagined feminist world like *Barbie*, or a temporary female space like the one in *Portrait of a Lady on Fire*. Her abusive marriage cannot be separated from migration, family responsibility, economic dependence, and the social systems around her. Her husband's violence is central, but he is not the only reason she cannot leave. The police and welfare system give no effective response; as an immigrant, she has no relatives or friends in Hong Kong that care about her... In this sense, the film makes domestic violence beyond a private conflict between one cruel husband and one suffering

wife; it's a violence created by the whole society (*Night and Fog*).

The way Hui films Xiaoling makes this even clearer. The camera stays close to the ordinary conditions of Xiaoling's life. We see her at home, taking care of her children, working, experiencing violence, and trying to find help. The tone is much darker than in the other two films. Xiaoling is angry, but she is also tired and afraid. Leaving is never shown as something she can easily decide to do. *Barbie*, an imaginary doll-lady, can eventually leave behind the role that has been designed for her; Marianne and Héloïse, at least for a short time, are able to create a space of their own, in part due to Héloïse's relative economic privilege, in part because as an unmarried, rural woman in the 17th century, she is kept in isolation. Xiaoling never really has such a space. Even her private life is shaped by lack of money, by marriage, children, migration, and the institutions she depends on. This is very close to the tension I discussed earlier in Hong Kong feminism: women may appear to have more choices in public life, while the family can still remain a place where those choices become much more limited, and women are actually given so little room to act (*Night and Fog*).

Taken together, these three films show that feminist filmmaking does not have one fixed cinematic language. *Barbie* makes gender politics visible through comedy, speeches, fantasy, and popular culture; *Portrait of a Lady on Fire* works through gaze, intimacy, silence, and memory; *Night and Fog* places women's choices inside the material realities of family, migration, poverty, and social institutions. Across different cultural contexts, filmmakers are finding very different ways to speak about female experience and challenge patriarchal structures. However, a film's meaning changes when it reaches audiences outside the cultural context of its creation. In the case of feminist film, the same feminist language may be accepted, resisted, simplified, or translated into new cultural terms.

As a Chinese filmmaker who hopes to continue making films about women and gender, I am especially interested in what happens when feminist cinematic languages developed in different cultural contexts meet Chinese audiences. Understanding these responses matters to me not only as a researcher, but also as a creator who may one day face the same question in my own work: how can I speak about women's experiences in ways that remain meaningful within the cultural context of the multiple audiences I hope to reach? Accordingly, the next section examines how these three films are interpreted when they enter Chinese cultural and media spaces.

4. Reception in China

4.1 Barbie

To investigate *Barbie*'s reception in China, I first examined comments under two highly engaged Bilibili videos about the film. Bilibili is one of China's most influential online platforms, especially among younger users—in 2025, the average age of its active users was 26—and is active in discussions of film, culture, and social issues. The first, titled "Using a Pink World to Deconstruct Patriarchy: *Barbie*'s Satire Goes Far Beyond This," had received more than 11.5 million views, 510,000 likes, and over 7,400 comments when I recorded the data. The second, "What If Your Boyfriend Is Offended by *Barbie*? What Is Hidden in Barbieland?", had received more than 570,000 views and over 8,300 comments. Though comments on these two videos cannot represent all Chinese viewers, they still provide useful gender-focused case studies of some common ways Chinese audiences understood and discussed *Barbie* (Bilibili Inc.; "Using a Pink World"; "What If Your Boyfriend").

Previous research shows that these Bilibili comments were not isolated examples. Jiao et al., who studied Chinese social-media responses to *Barbie*, found strong disagreements around feminism, with different gender groups and platforms responding to the film in different ways. Their study provides a broader background for my analysis (Jiao et al.).

One common response was to describe *Barbie* less as a "feminist" film and more as a film about "gender equality" (平权). One of the most-liked comments under the first video, with more than 44,000 likes, states that "what *Barbie* really wants to talk about is equality" (平权). Another commenter describes the film as "truly moderate equality" (真正的温和平权), adding that if even this view is considered "extreme," then other ideas about equality would probably not be accepted. In the Chinese online context, discussions of feminism are often connected with terms such as "gender equality" (平权), "extreme feminism" (极端女权), and "gender antagonism" (性别对立). These words are drawing boundaries around feminism: what level of feminism is "acceptable"? And when does it become "extreme"? Some Chinese viewers are more willing to accept feminist ideas framed as "gender equality" rather than "feminism," since feminists are often perceived in Chinese culture as being "extreme" or creating gender conflict ("Using a Pink World"; "What If Your Boyfriend").

A similar tendency of making the film's feminist message less gender-specific appears in another comment. The commenter argues that if "women" in Gloria's speech were

replaced with “men,” much of the speech would still make sense, since both men and women are expected to become successful, improve themselves, and appear “perfect.” They use the Chinese internet expression “摆烂” (*bailan*)—giving up on the pressure to keep competing and improving—as a response to these expectations. In these comments, Gloria’s speech is not just about pressures placed on women; it becomes a criticism of the pressure everyone in society faces. This framing easily connects with the experiences of many Chinese young people, but it avoids the gender-specific meaning of Gloria’s speech. This de-gendered perspective weakens concerns about why women bear these pressures and how these pressures are related to the gender structure (“Using a Pink World”; “What If Your Boyfriend”).

Like the earlier comments about “gender equality,” this interpretation turns *Barbie*’s feminism into a broader and less confrontational social message. This is a kind of reception with strong Chinese context: the audience translates feminist criticism into a more general social criticism that is less suggestive of gender conflict. Otherwise, the discussion can easily raise concerns about “gender antagonism” (性别对立). Guan and Shang (2026), for example, find that “gender antagonism” has become a common term on Chinese social media and, on Weibo, even a category used in platform governance. When both men and women are put into the same category of “the oppressed”, however, the power differences between genders would no longer be the center of discussion. The center of the problem shifts from power inequality between genders to shared suffering under social pressure (Guan and Shang; “Using a Pink World”).

The comments also translate *Barbie* into Chinese, the popular-cultural language of the "independent woman” narrative. One highly-liked comment calls *Barbie* “my first 大女主,”: 大女主 is Chinese term commonly used for a female protagonist whose story emphasizes her own agency and central position. Chinese audiences don’t directly translate the western theoretical expression of “female subjectivity” into Chinese, but rather put *Barbie* into a familiar local category to better understand it. This makes *Barbie*’s feminism more familiar to Chinese viewers, but also shifts attention away from structural gender inequality, reducing its broader social message to that of an individual woman becoming strong and independent. As a result, feminism becomes a story of individual self-improvement, while the social structures that produce gender inequality receive less attention (“Using a Pink World”).

Moreover, the discussion of *Barbie* moved beyond the film itself and into viewers’ everyday relationships. The title of the second Bilibili video asks what happens when a

"boyfriend" becomes offended while watching *Barbie*, suggesting that the film can reveal differences in how couples think about gender. This framing was not unique to that video. In July 2023, *Sixth Tone* reported that some Chinese social-media users used men's reactions to *Barbie* as a way to evaluate dates or boyfriends, and many others shared stories of couples arguing over the film's gender politics. In my Bilibili sample, one man commenter watched the film with his girlfriend, but his girlfriend found it boring and felt that it was somewhat unfair to men, while he enjoyed the film and thought it was "pretty good." He makes fun of his girlfriend and himself, as being "the opposite of other couples." A reply says that they both are great because she was able to empathize with men while he was able to empathize with women. These comments reveal *Barbie*'s function as a way for couples to discover and discuss differences in how they understand gender, and that these differences did not always follow the expected male-versus-female division. Additionally, another commenter describes first hearing about the film in a conversation with a friend who later posted their discussion on *Barbie* to WeChat Moments. These examples show that the film's reception did not stay inside cinemas or film-review spaces. Its ideas move through conversations, friendships, social-media feeds, and everyday relationships, becoming part of how Chinese people discuss gender in everyday life (Yang; "What If Your Boyfriend").

I also see many opposing voices among comments. Yet criticism and reinterpretation are part of the reception. One commenter writes, "*Barbie* still has many problems, but we need more films like *Barbie*." The importance of "we need more *Barbie*" suggests that mainstream films openly centered on women's experiences and gender inequality still feel relatively rare to some Chinese viewers. The film may be imperfect, and its feminism may seem too direct or even uncomfortable to some audiences, but this discomfort is part of the reasons that make such films valuable in China: by generating so much debate around "gender equality," "extreme feminism," "gender antagonism," and the "大女主" (independent female protagonist) they bring feminist questions into a mainstream cultural space where these questions are still contested. Those of us who feel this way are not hoping that one film can settle these debates, but that more films like *Barbie* can make conversations about women and gender increasingly visible and difficult to dismiss ("Using a Pink World"; "What If Your Boyfriend").

Such conversations develop beyond individual reactions. Yuan et al., based on interviews with 23 Chinese feminist fans of *Barbie*, found that participants built forms of collective identity around the film and adapted *Barbie* into a symbol of empowerment

connected to Chinese social realities. In this sense, the value of *Barbie* in China may manifest not only in the values it spreads, but also in the conversations it allows to continue and the community that it creates for ongoing gender discussions (Yuan et al.).

4.2 Portrait of a Lady on Fire

In my following investigation of *Portrait of a Lady on Fire*'s reception in China, I use a different set of materials from those I used for *Barbie*. In place of using comment sections under popular videos online, I examine the film's exhibition at the Beijing International Film Festival, long-form reviews on Douban and Bilibili, and discussions by Chinese film critics and scholars. This difference partly reflects how the two films entered Chinese culture: *Barbie* arrived as a mainstream commercial event and generated large public arguments about feminism, while *Portrait of a Lady on Fire* received less attention, mainly circulating through film-festival, cinephile, feminist, and queer film spaces. Obviously, these materials cannot represent all Chinese viewers, but they help us get a glance into how the film was understood within Chinese feminist and art-house film culture.

The film was introduced to Chinese audiences through a female-centered context. At the 2020 Beijing International Film Festival, *Portrait of a Lady on Fire* was selected for the newly created "Women's Voices" (女性之声) section, which focused on women's lives and inner experiences. It received five screenings. Chinese director Chen Zhengdao, recommending the film for the festival, described it as a rare film centered entirely on women's love. He also praised the delicacy of its emotions and visual style. Therefore, the film was introduced to Chinese audiences not simply as a romance, but as a film centered on women and women's experiences (Ni; Zhou).

One word appears particularly often in Chinese discussions of the film: equality (平等). Chinese feminist film scholar Dai Jinhua's discussion of the film is titled around the idea of a passionate and equal gaze, and she argues that the film uses a same-sex love story to imagine the possibility of equality in intimate relationships. For Dai, this equality happens partly because the relationship is between two women: unlike heterosexual relationships that already exist within a gender hierarchy, Marianne and Héloïse do not begin with the same fixed structure of male power and female submission. A Bilibili video essay is titled "The Gravity of Equality" (平等的重力), while another video centers around the question "Can equal love exist?" A Douban reviewer writing about a Beijing International Film Festival screening also uses the phrase "mutual and equal gaze" (双向且平等的凝视) to describe

Marianne and Héloïse's relationship. These titles show how strongly the idea of equality shapes Chinese discussions of the film (Dai; "The Gravity of Equality"; "Can Equal Love Exist?"; 灵魂漫谈录).

It's important to make it clear the definition of "equality" in these discussions. When commenting on this film, Chinese audiences not only use the term as a general statement that women should be equal to men. They are also describing a relationship in which neither woman remains fixed as the more powerful subject or the passive object. In this sense, "equality" refers not only to power between genders, but also more broadly to how power operates within an intimate relationship (Dai; 灵魂漫谈录).

However, this audience focus on intimate relationships does not mean that gender becomes unimportant. Within the terms of the film, gender is central to this equality, and to why it matters. Héloïse is initially being turned, by a male stranger, into a portrait for another male stranger who will decide whether to marry her. She is placed in the position of being looked at and judged by men. Marianne's relationship with her gradually changes this structure: Héloïse looks back, criticizes the portrait, studies Marianne in return, and participates in how she is represented. Their equality does not come simply from the fact that they are two women. It comes from the film creating a temporary relationship outside the usual structures of heterosexual marriage and the male artist–female muse tradition, where the roles of subject and object can be renegotiated (*Portrait of a Lady on Fire*).

This question of the discussion of "equality" is not uniquely Chinese. What is more specific to the Chinese reception I examined is that critics prefer to use the language of "equality" (平等), "mutual gaze" (互相凝视), and "equal gaze" (平等凝视) to explain this relationship. "Gender equality", which is more familiar and less contested to Chinese people, is used more often in contemporary Chinese gender discourse than "feminism", which is often considered as a Westernized, elitist, or controversial concept. Yet here, "equality" does not make feminism sound more moderate, as it sometimes does in the reception of *Barbie*. Instead, it becomes a way to describe how power works inside the relationship between Marianne and Héloïse: Neither person remains permanently the subject, and neither is fixed as the object (Dai; Jiao et al.).

This can be seen clearly in one long Douban review, which argues that Sciamma changes the traditional power structure of "gazing" (凝视) into a more attentive action of "looking" (注视). Marianne observes Héloïse in order to paint her, but Héloïse also looks back, studies Marianne, and later points out Marianne's own habits and gestures. The painter

watches the model, but the model also watches the painter. In the reviewer's interpretation, looking becomes something created between two subjects rather than a one-directional act in which one person controls the other (voidtakowlands).

"Female gaze" is thus translated into Chinese discussions in the following way: not simply that a woman replaces a man to become the person who looks, but that the person being looked at is allowed to look back and participate in how she is seen. Dai Jinhua argues that the traditional relationship between "artist" and "muse" is changed in this film because the gaze no longer belongs entirely to the artist. In this view, the female gaze is not a reversal of power, but a way of reciprocal looking (Dai).

Some Chinese critics further broaden the film's idea of equality beyond its specific gender context. Marianne and Héloïse's relationship becomes a way to ask whether any intimate relationship can avoid a fixed hierarchy between a more powerful subject and a more passive partner. This makes the film easier to connect with relationships in general, but it can also move attention away from the gender conditions that make this equality possible in the first place (Dai; "Can Equal Love Exist?").

This is different from the de-gendering I found in the Chinese reception of *Barbie*. In *Barbie*, localizing feminism through "equality" can sometimes weaken its gender-specific criticism by turning women's experiences into a more general story about pressure faced by everyone in society. In *Portrait*, broadening "equality" beyond gender instead turns a specifically female and queer relationship into a larger question about equality in human relationships. In both cases, gender-specific meanings become more universal, but they do so in different ways (Dai; Jiao et al.).

Another notable pattern is that many Chinese discussions do not enter the film through feminist theory. Instead, viewers often lead by talking about its beauty and emotions. Many Douban reviews mention its oil-painting-like images, restrained emotions, silence, memory, the Orpheus myth, and the final scene with Vivaldi. And many also focus on the plot, such as the sadness of the love story and the memory left after the two women separate (灵魂漫谈录; BaladeurS).

Since *Portrait of a Lady on Fire* is also received in China more as an art-house love story than a feminist or queer film, its impact is very different from *Barbie*'s. *Barbie* directly asks its audience to respond to gender issues and expresses direct criticism of patriarchy. In *Portrait*, on the contrary, feminism is expressed indirectly through scenes of daily life, through a mutual gaze. Attracted by the film's visual style or emotions, some Chinese

viewers focus much more on visual style and emotion than on the film's gender politics (灵魂漫谈录; BaladeurS).

Moreover, the reception of the film changes with personal experience. One Douban writer notes that she disliked the film when first watching it in university, but when watching it again several years later after a breakup, she responded to it very differently. Another reviewer also says that rewatching the film let them recall those emotional details that they had missed before. Though not representative of all Chinese audiences, these reviews still suggest another difference from the reception of *Barbie*. *Barbie* brings an immediate reaction: viewers quickly debate its feminism, while *Portrait* creates a slower and more personal kind of reception, in which the meaning of the film changes with audiences' own memories, relationships, and rewatching (都市丽踪; BaladeurS).

This also helps explain why discussions of *Portrait* have continued in Chinese online film culture years after its release. Bilibili video essays still return to themes such as equality, looking and being looked at, restraint, and memory. Rather than creating one large moment of public debate, the film continues to be revisited and reinterpreted within smaller cinephile and feminist film communities ("The Gravity of Equality"; "The Infinity of Restraint"; "Can Equal Love Exist?").

One experience we can draw from this comparison of the two films' reception is that the cultural differences that shape cross-cultural reception do not necessarily weaken or strengthen the feminist message of the film as originally understood in its own cultural context. In the reception of *Barbie*, translating feminism into "equality" makes its gender-specific criticism appear more general and less confrontational. Conversely, in *Portrait*, "equality" helps Chinese viewers describe a different structure of looking and loving. Meanwhile, approaching the film mostly through beauty and emotion can make its feminism easier to enter, but may also move attention away from its specific queer and political context. To conclude, Chinese audiences are not just accepting or rejecting the film's feminist ideas. They are selecting different parts of the film and connecting them with their own ways of understanding relationships, art, gender, and memory.

4.3 Night and Fog

My investigation of the reception of *Night and Fog* (2009, Ann Hui) in China is slightly different from that of the previous two films. Unlike *Barbie* and *Portrait of a Lady on Fire*, *Night and Fog* is not a "Western feminist film entering China," but rather a Chinese

social story being reinterpreted across different Chinese cultural contexts. As a Chinese-language Hong Kong film, *Night and Fog* is based on a real domestic-violence case in Tin Shui Wai. Therefore, I looked at how this story was understood across different Chinese spaces, including mainland film reviews, Bilibili video essays, discussions of Hong Kong society, and interviews with Ann Hui. Compared with the previous two films, its reception is less centered on feminism. Beyond gender, viewers also discuss poverty, migration, family, social welfare, or the failures of public institutions. This complexity makes the feminist meaning of *Night and Fog* only one part of a broader discussion about social inequality and marginalization (*Night and Fog*; "许鞍华｜谈电影").

Interestingly, one of the most visible points of entry into mainland online platforms is through discussions of domestic violence (家暴). Bilibili videos use titles such as "more terrifying than a horror film" and "explaining domestic violence completely." One of these videos introduces the film through the "2004 family tragedy" adding that domestic violence either happens "zero times or countless times." Many videos place the film under categories such as crime and domestic violence ("More Terrifying than a Horror Film").

These labels are eye-catching, but they narrow the way viewers approach the film. When *Night and Fog* is mainly described as a horrifying story of domestic violence or a real-life family tragedy, attention is easily drawn to how violent Li Sen is and how shocking the case is. Viewers may also focus too much on Xiaoling's visible fear, suffering, or even appearance; some may even "watch" her pain with a kind of prurient curiosity. Yet the film spends much more time asking why Xiaoling remains trapped even after she tries to leave and seek help. Poverty, migration, economic dependence, social isolation, and failures of social support are all part of that tragedy. We cannot simplify the film or blur its true story. This type of propaganda that shifts the focus shows how platforms can influence reception. A title such as "more terrifying than a horror film" can attract viewers quickly. In this sense, reception is shaped not only by what viewers say after watching a film, but also by how the film is packaged before they watch it ("More Terrifying than a Horror Film"; *Night and Fog*).

More detailed Chinese discussions often expand the story beyond domestic violence. Xiaoling has many identities. Besides being a woman abused by her husband, she is also a mainland woman who moves to Hong Kong through marriage, lives in a low-income community, has limited local connections. She depends on a husband who has more social and economic power. Mainland Chinese film reviews and media discussions of Tin Shui Wai understand the film in terms of the difficulties faced by new mainland immigrants and low-

income families in Hong Kong. Contemporary reporting noted that domestic violence was a serious issue among some new-immigrant families and directly mentioned *Night and Fog* when discussing this problem. Xiaoling's vulnerability cannot just be explained through gender. It is intersectional, connected with class, migration, and regional inequality altogether (炫风 and 曾明瑞).

This leads to one of the most important questions in the film's reception. Many people ask, "Why didn't she leave?" But Xiaoling does try to leave. Both the film and the real case show that Xiaoling sought help many times and moved toward a more independent life, yet these attempts did not allow her to escape the violence. Female "agency" here is not merely a choice, but heavily depends on whether the conditions around her make that choice possible (*Night and Fog*; "许鞍华｜谈电影").

The perception of female "agency" in *Night and Fog* poses an interesting contrast with the reception of *Barbie*. As mentioned above, some Chinese viewers translate *Barbie*'s feminism into "大女主": a woman becoming independent, strong, and in control of her own life. While *Night and Fog* shows the limits of that idea, some viewers nonetheless struggle to perceive those limits.

Another difficult question centers on Li Sen, the husband. The film does not present him as a completely unexplained "monster." His unemployment, economic frustration, aging, insecurity, and need for control are all shown as part of the background of his violence. These details suggest that domestic violence does not appear from nowhere; it can also be connected with economic pressure, masculinity, and a loss of social status (*Night and Fog*).

However, explaining these conditions creates a risk. If viewers spend too much time asking why Li Sen became violent, they may begin to sympathize with his difficulties and shift their attention away from his own responsibility and Xiaoling's suffering. Yet if he is treated only as an evil individual, the larger social conditions behind the violence disappear again. This creates a difficult balance: understanding the social causes of violence is necessary, but explanation should not become an excuse for the person who commits it, nor should it lead, as it does in Chinese reception of *Barbie*, to false equivalences (*Night and Fog*).

Overall, the Chinese reception of *Night and Fog* shows that gender violence cannot be understood through gender alone. Depending on what viewers focus on, the same film can become a shocking domestic-violence case, a story about a mainland migrant woman, a criticism of social welfare, or a discussion of poverty and unequal marriage. Each

interpretation makes one part of the film more visible, but may also push other parts into the background.

5. Conclusion

The reception of these three films in China shows that feminist ideas change when they travel across cultures. In *Barbie*, many Chinese viewers translate feminism into familiar terms such as "gender equality" (平权) and the "大女主," sometimes making its criticism of gender inequality more general and less confrontational. In *Portrait of a Lady on Fire*, the same language of "equality" works differently: it helps viewers describe mutual gaze, equal love, and a relationship in which neither woman is reduced to an object. The film is also often approached through beauty, emotion, and memory, which make its feminist and queer meanings less visible. For *Night and Fog*, while online discussions often introduce the film through domestic violence and a shocking real case, more detailed readings connect the woman's situation to migration, poverty, family, and failures of social support. Across the three films, Chinese audiences emphasize different aspects of feminism depending on the cinematic form of the film and the cultural language available to them. Film is not only a way to transmit feminist ideas, but also a space where those ideas are changed through contact with different audiences.

Seen together, these films suggest that there is no single global feminist conversation moving in one direction. Instead, that conversation grows through difference. American, French, Hong Kong, and Chinese cultural contexts bring different histories, priorities, and vocabularies to questions about women's freedom, desire, representation, and power. When feminist films move between these contexts, some meanings are softened, others become more visible, and new questions emerge. These changes allow the conversation to morph, deepen, and continue, as viewers and filmmakers use different cultural experiences to reconsider what equality, agency, and female subjectivity can mean.

Bibliography

Alexander, Kerri Lee. "Feminism: The First Wave." *National Women's History Museum*, 5 Apr. 2021, https://www.womenshistory.org/exhibits/feminism-first-wave-0. Accessed 26 Oct. 2026.

All About Love. Directed by Ann Hui, 2010.

BaladeurS. "假如重逢 [If They Met Again]." *Douban*, 20 Apr. 2020, https://movie.douban.com/review/12524131/. Accessed 26 Oct. 2026.

The Balconettes. Directed by Noémie Merlant, 2024.

Barbie. Directed by Greta Gerwig, Warner Bros., 2023.

Beauvoir, Simone de. *The Second Sex*. Translated by Constance Borde and Sheila Malovany-Chevallier, Vintage, 2011.

Bilibili Inc. *Q2 2025 Bilibili Inc. Investor Presentation*. 2025, https://ir.bilibili.com/media/ayhjclou/q2-2025-bilibili-inc-investor-presentation.pdf. Accessed 26 Oct. 2026.

The Bride!. Directed by Maggie Gyllenhaal, Warner Bros., 2026.

Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge, 1990.

"Can Equal Love Exist? [银河影评：燃烧女子的肖像-上：平等的爱有可能存在吗]." *Bilibili*, 17 Dec. 2022, https://www.bilibili.com/video/BV12V4y1A718/. Accessed 26 Oct. 2026.

China Film Administration. "电影产业促进法 [Film Industry Promotion Law]." *China Film Administration*, 7 Nov. 2016, https://www.chinafilm.gov.cn/xxgk/zcfg/flfg/201611/t20161107_1460.html. Accessed 26 Oct. 2026.

Cixous, Hélène. "The Laugh of the Medusa." Translated by Keith Cohen and Paula Cohen, *Signs*, vol. 1, no. 4, 1976, pp. 875–93. https://doi.org/10.1086/493306.

Cléo from 5 to 7. Directed by Agnès Varda, 1962.

Crenshaw, Kimberlé. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum*, 1989, pp. 139–67.

Dai, Jinhua. "《燃烧女子的肖像》：炙热的凝视，平等的凝视 [Portrait of a Lady on Fire: A Passionate Gaze, an Equal Gaze]." *The Paper*, 15 Feb. 2022, https://www.thepaper.cn/newsDetail_forward_16694166. Accessed 26 Oct. 2026.

Dargis, Manohla. "'The Bride!' Review: Frankie, My Dear." *The New York Times*, 5 Mar. 2026, https://www.nytimes.com/2026/03/05/movies/the-bride-review.html. Accessed 26 Oct. 2026.

Gonzalez, Umberto. "Female Directors See 'Complete Reversal' of Progress as 2025 Marks 5% Drop in Film Work, Study Finds." *TheWrap*, 31 Dec. 2025, https://www.thewrap.com/creative-content/movies/female-directors-work-drops-5-points-2025/. Accessed 26 Oct. 2026.

"The Gravity of Equality [《燃烧女子的肖像》评析—平等的重力]." *Bilibili*, uploaded by 锈山, 21 Mar. 2020, https://www.bilibili.com/video/BV1KE411A7oc/. Accessed 26 Oct. 2026.

Guan, Zexu, and Zheyu Shang. "Constructing Gender Antagonism: Moralised Platform Governance in China." *Media and Communication*, vol. 14, 2026, article 12064. https://doi.org/10.17645/mac.12064.

Half the Picture. Directed by Amy Adrion, 2018.

hooks, bell. *Feminist Theory: From Margin to Center*. South End Press, 1984.

"The Infinity of Restraint [《燃烧女子的肖像》评析—克制的无穷]." *Bilibili*, uploaded by 锈山, 17 Apr. 2020, https://www.bilibili.com/video/BV1Ta4y147KR/. Accessed 26 Oct. 2026.

Irigaray, Luce. *This Sex Which Is Not One*. Translated by Catherine Porter with Carolyn Burke, Cornell UP, 1985.

Jiao, Lin, et al. "Polarized Views over Feminism: Gendered Reception of Barbie among Chinese Audiences on Social Media." *Feminist Media Studies*, vol. 26, no. 5, 2026, pp. 1581–98. https://doi.org/10.1080/14680777.2025.2478390.

Kristeva, Julia. *Revolution in Poetic Language*. Translated by Margaret Waller, Columbia UP, 1984.

Lee, Eliza Wing-Yee, editor. *Gender and Change in Hong Kong: Globalization, Postcolonialism, and Chinese Patriarchy*. UBC Press, 2004.

Library of Congress. "Ruth Handler, Entrepreneur, Introduces the Barbie Doll." *This Month in Business History*, https://guides.loc.gov/this-month-in-business-history/march/ruth-handler-introduces-barbie. Accessed 26 Oct. 2026.

Library of Congress. "20th Century: The Waves of French Feminism." *French Women & Feminists in History: A Resource Guide*, https://guides.loc.gov/feminism-french-women-history/20th-century. Accessed 26 Oct. 2026.

Lim, Adelyn. *Transnational Feminism and Women's Movements in Post-1997 Hong Kong: Solidarity Beyond the State*. Hong Kong UP, 2015. https://doi.org/10.5790/hongkong/9789888139378.001.0001.

Little Women. Directed by Greta Gerwig, Sony Pictures Releasing, 2019.

Lorde, Audre. *Sister Outsider: Essays and Speeches*. Crossing Press, 1984.

"More Terrifying than a Horror Film: Explaining Domestic Violence [〖越哥〗让人头皮发麻的华语电影，比恐怖片还恐怖，说透了家庭暴力这件事！《天水围的夜与雾》]." *Bilibili*, uploaded by 越哥说电影, 12 June 2020, https://www.bilibili.com/video/BV1UK4y1x7AP/. Accessed 26 Oct. 2026.

Mulvey, Laura. "Visual Pleasure and Narrative Cinema." *Screen*, vol. 16, no. 3, 1975, pp. 6–18. https://doi.org/10.1093/screen/16.3.6.

Munro, Ealasaid. "Feminism: A Fourth Wave?" *Political Insight*, vol. 4, no. 2, 2013, pp. 22–25. https://doi.org/10.1111/2041-9066.12021.

National Archives. "19th Amendment to the U.S. Constitution: Women's Right to Vote (1920)." *National Archives*, https://www.archives.gov/milestone-documents/19th-amendment. Accessed 26 Oct. 2026.

News from Home. Directed by Chantal Akerman, 1976.

Ni, Wei. "北影节展映将开启，侯麦影展、《82 年生的金智英》等登陆影院 [Beijing International Film Festival Screenings to Begin]." *The Beijing News*, 13 Aug. 2020, https://m.bjnews.com.cn/detail/159728777815963.html. Accessed 26 Oct. 2026.

Night and Fog. Directed by Ann Hui, 2009.

Petite Maman. Directed by Céline Sciamma, 2021.

Portrait of a Lady on Fire. Directed by Céline Sciamma, 2019.

Promising Young Woman. Directed by Emerald Fennell, Focus Features, 2020.

Shaw, Gareth, and Xiaoling Zhang. "Cyberspace and Gay Rights in a Digital China: Queer Documentary Filmmaking under State Censorship." *China Information*, vol. 32, no. 2, 2018, pp. 270–92. https://doi.org/10.1177/0920203X17734134.

Smith, Stacy L., et al. *Inclusion in the Director's Chair? Gender, Race, & Age of Film Directors Across 1,000 Films from 2007 to 2016*. USC Annenberg Media, Diversity, & Social Change Initiative, 2017, https://annenberg.usc.edu/sites/default/files/2017/04/06/MDSCI_Inclusion%20_in_the_Directors_Chair.pdf. Accessed 26 Oct. 2026.

Stevens, Isabel. "No Man's Land: Céline Sciamma on Portrait of a Lady on Fire." *Sight and*

Sound, British Film Institute, 8 Dec. 2022, https://www.bfi.org.uk/sight-and-sound/features/no-mans-land-celine-sciamma-portrait-lady-fire. Accessed 26 Oct. 2026.

Sun, Haocan, and Jinghong Xu. "Gender Stereotypes and Composition in the Chinese Film Industry: A Study of the Top 15 Grossing Films." *Journal of Broadcasting & Electronic Media*, vol. 68, no. 1, 2024, pp. 125–44. https://doi.org/10.1080/08838151.2023.2298276.

University of Amsterdam. "Chinese, Female Filmmakers through a Feminist Lens." 6 Jan. 2022, https://www.uva.nl/en/shared-content/faculteiten/en/faculteit-der-geesteswetenschappen/news/2022/01/chinese-female-filmmakers-through-a-feminist-lens.html. Accessed 26 Oct. 2026.

"Using a Pink World to Deconstruct Patriarchy: Barbie's Satire Goes Far Beyond This [用粉色世界瓦解父权制度｜《芭比》的讽刺，远不止这些]." *Bilibili*, uploaded by 平原小狐, 24 July 2023, https://www.bilibili.com/video/BV1n8411m78H/. Accessed 3 Oct. 2026.

voidtakowlands. "浅析《燃烧女子的肖像》中的目光 [A Brief Analysis of the Gaze in Portrait of a Lady on Fire]." *Douban*, 2 July 2022, https://movie.douban.com/review/14489951/. Accessed 26 Oct. 2026.

Water Lilies. Directed by Céline Sciamma, 2007.

The Way We Are. Directed by Ann Hui, 2008.

"What If Your Boyfriend Gets Defensive Watching Barbie? What Does Barbieland Symbolize? [男朋友看《芭比》破防了怎么办？芭比世界有什么隐喻？]." *Bilibili*, uploaded by 沈奕斐, 2 Aug. 2023, https://www.bilibili.com/video/BV1AF4117tfF/. Accessed 3 Oct. 2026.

Yang, Caini. "Despite Slow China Start, 'Barbie' Sets a Benchmark — For Dating." *Sixth Tone*, 24 July 2023, https://www.sixthtone.com/news/1013374. Accessed 26 Oct. 2026.

Yuan, Xinyao, et al. "Feminist Fan Engagement and Collective Agency: A Case Study of Barbie in China." *International Journal of Cultural Studies*, vol. 29, no. 2, 2026, pp. 291–310. https://doi.org/10.1177/13678779251358881.

Zhou, Huixiaowan. "北影节 300 部佳片，李少红、李晨、黄晓明等推荐这 20 部 [Twenty Recommended Films at the Beijing International Film Festival]." *The Beijing News*, 20 Aug. 2020, https://m.bjnews.com.cn/detail/159790578015295.html. Accessed 26

Oct. 2026.

灵魂漫谈录. “北影节女性之声单元：《燃烧女子的肖像》 [BJIFF Women’s Voices: Portrait of a Lady on Fire].” *Douban*, 2 Sept. 2020, https://movie.douban.com/review/12843965/. Accessed 26 Oct. 2026.

炫风, and 曾明瑞. “‘悲情市镇’天水围 [Tin Shui Wai, a ‘Town of Sorrow’].” *南都周刊 [Southern Metropolis Weekly]*, 15 July 2011, https://news.sina.com.cn/c/sd/2011-07-15/152122820895_2.shtml. Accessed 26 Oct. 2026.

“许鞍华｜谈电影，谈上海，谈《第一炉香》 [Ann Hui on Film, Shanghai, and Love after Love].” *The Paper*, 22 Oct. 2021, https://www.thepaper.cn/newsDetail_forward_15009422. Accessed 26 Oct. 2026.

都市丽踢. “五年后二刷，才看懂 [I Understood It Only after Rewatching Five Years Later].” *Douban*, 5 Sept. 2025, https://movie.douban.com/review/17041570/. Accessed 26 Oct. 2026.