

PROMETHEUS

2022 Edition



The Johns Hopkins University
Undergraduate Journal of Philosophy

Gillan Chalono
Long Nguyen
Sabah Rainey
Tobias Rivas
Weiming Sheng

The Personal is Political: Property as a Distributive Interface
Language and Common Sense in Berkeley's Philosophy
New Sights and Signs: Street art as a Response to Willful Hermeneutical Ignorance
Reasoning Together
Plato's Tripartite Classification of Normative Reasons

Prometheus Undergraduate Philosophy Club and Journal Academic Year 2021 – 2022

President	Lawrence Cao
Advisor	Cara Cummings
Undergraduate Editors	Zayeed Akhter, Shour Arashanapalli, Damon Choi, Jacky Hu, Clara Kraft, Ava Lasmanis, Demi Liu, Cecilia Miller, Madeleine Schmitz, Nhat Tran, Audrey Zeldin, Zhaojun Zhai
Graduate Student Help	Chris Arledge, Austin Heath, Maegan Kaczmarek, Jacob Lettie, Tammo Lossau, Itai Marom, Ryan Ross.

The Johns Hopkins University
Baltimore, Maryland

Copyright © 2022 Prometheus
All rights reserved.

A Note from the President of Prometheus

Dear Readers,

On behalf of the Prometheus Editorial Board, I am excited to present the Volume 14 of *Prometheus: The Johns Hopkins University Journal of Philosophy*. *Prometheus* is a philosophy journal established in 2009 by undergraduate students at Johns Hopkins University. Based on our web analytics, the 2021 edition of Prometheus has already reached a broad audience, recording 500 views or downloads within the first eight months of its publication. Starting this year, Prometheus will replace our year-edition catalog system with the standard volume system. Since the first issue of Prometheus was published in 2009, the 2022 edition of Prometheus will be the 14th volume.

We are highly impressed by the quality of this year's submissions. Thank you for everyone who submitted a paper. We also wish to thank the Johns Hopkins University Alumni Association and the William H. Miller III Department of Philosophy for their generous support. In addition, we are grateful to everyone who has helped Prometheus, especially Michelle Brock, Rebecca Swisdak, and Professor Steven Gross, for their administrative support, our advisor Cara Cummings for her teaching and advice, our editors for their hard work, and our board members leadership and excellent service.

Lawrence Cao

Baltimore, May 2022

The Personal is Political: Property as a Distributive Interface

GILLAN CHALONO

AMHERST COLLEGE

Abstract: In “Where the Action Is: On the Site of Distributive Justice,” G.A. Cohen provides a scalding critique of John Rawls’s *A Theory of Justice* by illuminating its shortcomings in addressing the distributive consequences of personal behavior. Rawls’s solution to achieving distributive justice— the ethical distribution of rights and resources in society— is to apply regulatory principles to major political and economic institutions. Cohen correctly identifies that a genuine concern for distributive justice should account for social practices that perpetuate unjust distributions of wealth and power within our private lives. Yet, out of respect for personal liberty, Rawls has proven remarkably reluctant to endorsing the regulation of interpersonal discrimination. My paper will respond to Cohen and provide a defense of Rawls by exploring the far-reaching effects of the legal institution of property. I will argue that property, while enforced through legal institutions, can still have a demonstrable impact on the distributive consequences of personal behavior. Since private associations and social practices do not occur outside of a legal context, I believe that property law offers Rawls an avenue to exert a corrective influence on discriminatory social practices that pose a problem for distributive justice without compromising his commitment to protecting basic liberties.

Keywords: John Rawls, Distributive Justice, G.A. Cohen, Law, Feminist Philosophy

I. Introduction

In *A Theory of Justice*, John Rawls presents his arguments for why applying ethical rules and principles to the basic structure of society is sufficient to produce a just distribution of rights and advantages¹. Careful of encroaching on personal freedom, Rawls generally defines the basic structure to include major political and economic institutions like the state and the free market, and to exclude the choices and associations that citizens make in their private life.

G.A. Cohen’s critique of Rawls’s basic structure, “Where the Action Is: On the Site of Distributive

Justice,” posits that Rawls’s reluctance to consider the effects of personal behavior is incongruent with a genuine concern for distributive justice². Personal behavior is often motivated by social conventions that discriminate on the basis of identity, leading to forms of social coercion that place unjust economic and social burdens on different members of society. As such, a basic structure that neglects personal behavior seems remarkably incomplete. But if Rawls does assent to regulating personal choices and associations, not only would he violate his commitment to basic liberty, but the notion of a “basic structure” would be rendered inapplicable since everything outside of the structure (e.g., personal choice and behavior) is now eligible for regulation. Cohen concludes that Rawls has no escape from this dilemma, so he ought to abandon his basic structure altogether.

In this paper, I will respond to Cohen and provide a defense of Rawls’s basic structure by exploring one of its principal institutions: property. I will argue that property, while enforced through our legal institutions, has a demonstrable impact on personal behavior and social convention. Since private associations and social practices do not occur outside of a legal context, I believe that property law can exert a corrective influence on the discriminatory social practices that pose a problem for distributive justice. In this way, the legal institution of property can offer Rawls an avenue to address the distributive consequences of personal behavior without compromising his commitment to basic liberties.

I will begin by briefly summarizing Rawls’s position on the basic structure. I will then consider Cohen’s arguments and demonstrate that a Rawlsian society could significantly mitigate the unjust distributive impact of sexist social practices through property legislation. By illuminating the relationship between property law and personal behavior, I will show how property can act as a distributive interface between the legal and the social, vindicating Rawls’s basic structure and elucidating the full import of the claim that the personal is political.

II. The Basic Structure

¹ See John Rawls, *A Theory of Justice* (Cambridge, MA: Belknap Press, 1999), hereafter referred to as *Theory*.

² See Cohen, G. A. “Where the Action Is: On the Site of Distributive Justice.” *Philosophy and Public Affairs*, vol. 26, no. 1, Dec. 1997, hereafter referred to as *Action*.

Rawls defines the basic structure to include the major political, economic, and social institutions of society. Since institutions like the state or the free market are primarily responsible for the societal distribution of rights and advantages, Rawls believes that principles of justice should regulate these institutions to eliminate unjust inequalities and produce fair prospects for all members of society³.

Alternatively, Rawls demonstrates a distinctive opposition to regulating personal liberties and freedoms. He maintains that his principles are irrelevant to “voluntary cooperative arrangements” and “private associations,” and his first principle, the principle of basic liberties, ensures that citizens are entitled to “the most extensive scheme of equal basic liberties” that is compatible with the liberties of others⁴. Rawls’s commitment to political and personal liberties, like the right to vote or freedom of speech, stems from his belief that all individuals should have the freedom to “pursue whatever plan of life” they choose for their own conception of happiness⁵. As such, Rawls agrees that social institutions like the “monogamous family” need legal protection because of their determinative influence on life outcomes, yet he remains unconcerned with regulating the quotidian choices that occur within them⁶.

To illustrate Rawls’s position on regulating personal behavior, consider the diverging application of his principles onto the institutions of the state and the family. His second principle, the difference principle, states that inequalities of wealth and power are permissible if and only if they are of benefit to the least well-off⁷. In a society with vast wealth inequality, the difference principle may require the state to enact some form of wealth redistribution to ameliorate conditions for the lowest earning members of society. In contrast, the difference principle would place virtually no responsibility onto the family. Rawls’s opposition to the regulation of private associations would bar any form of legal intervention into familial relationships regardless of how these relationships may perpetuate unjust inequalities. To Rawls, basic liberties like the right to

free association cannot be forfeited for “greater social and economic advantages”⁸.

III. Social Convention and the Basic Structure

Dilemma

Cohen’s critique of Rawls is informed by the feminist maxim that the personal is political⁹. This maxim reflects that unjust gender discrimination occurs in personal contexts outside of the traditional political arena. Instead of continuing down this line of argumentation, Cohen attempts to abstract away from the gendered content of this maxim by arguing that Rawls’s basic structure fails to acknowledge that personal “choices not regulated by the law fall within the primary purview of justice”¹⁰. He contends that the basic structure is incapable of addressing the distributive consequences of social practices in general.

Cohen attempts to prove this incapability by differentiating between legal and social coercion. According to Cohen, Rawls’s basic structure is widely interpreted as composed of legally coercive institutions¹¹. These institutions use the law to determine “what people may and must do”¹². Consider the legal obligation imposed by the state onto automobile drivers to require the use of seatbelts, or the legal consequences attached to mandatory attendance in primary schools. These legal mandates coerce citizens to adhere to certain types of behavior. Alternatively, social coercion is generated by institutions whose rules are primarily formed by “convention, usage, and expectations”¹³. Unlike legal institutions, socially coercive institutions have no formal constitutions where rules can be explicitly enumerated. Rather, their rules and expectations are constituted and enforced exclusively through *behavior*.

An example of a socially coercive institution is the heteronormative family. There is no law that forces women to disproportionately perform domestic labor or to not participate in the workforce. Instead, these expectations are created and enforced through patterns of legally protected behavior. Such behavior may include

³ *Theory*, pp. 6-10

⁴ *Theory*, pp. 52-55, see also p. 7

⁵ *Theory*, pp. 80-81

⁶ *Theory*, pp. 6-7

⁷ See *Theory*, “Interpretations of the Second Principle,” pp. 57-65

⁸ *Theory*, pp. 53-54

⁹ Cohen’s critique largely drew its feminist inspiration from Okin, Susan Moller. 1987. “Justice and Gender.” *Philosophy & Public Affairs* 16 (1): 42-72.

¹⁰ *Action*, p. 4

¹¹ Rawls himself affirms the consistency of this interpretation in *Political Liberalism*, pp. 227-30

¹² *Action*, p. 19

¹³ *Action*, pp. 19-20

encouraging children to participate in gender-stereotypical activities, exclusively representing women as caretakers in media, or censoring women in public debates. Although these behaviors are not legally coercive, they create social expectations that women should carry greater domestic burdens solely due to a morally arbitrary characteristic like gender.

Cohen differentiates between legal and social compulsion to argue that Rawls's basic structure is victim to an irreconcilable dilemma. Since social coercion can result in unjust distributions of advantages and burdens, Cohen posits that a basic structure restricted to legal coercion cannot adequately address concerns of distributive justice. But directly regulating social coercion is not an avenue open to Rawls. His commitment to basic liberties forbids the regulation of socially coercive behavior. Furthermore, even if Rawls did acquiesce to the regulation of personal behavior, his identification of a societal basic structure would cease to draw any conceptual utility. Personal behavior is antithetical to the notion of a structure or an institution; one could hardly say that a senator's choice to drink wine instead of champagne is necessarily a state decision solely because she occupies a government office. Ultimately, this dilemma leads Cohen to conclude that Rawls's basic structure is incapable of adequately producing distributive justice in a society riddled with unjust social coercion¹⁴.

IV. The Institution of Property

My defense of Rawls rests on what I believe to be an ironic flaw in Cohen's reasoning. By dividing coercion into the categories of "legal" and "social," Cohen reaffirms the false separation between the personal and the political. Social coercion does not exist as a phenomenon wholly segregated from legal contexts, and legal institutions act as the matrix in which our social practices occur. Therefore, I believe that property offers Rawls a way out of the basic structure dilemma. Instead of admitting social practices into the basic structure and forgoing his commitment to basic liberties, Rawls can enact changes to legally coercive institutions that will greatly limit the ease with which

we can produce unjust distributions within our personal lives.

The legally coercive institution I will focus on in this paper will be the institution of property. By property, I loosely refer to the rights, entitlements, and obligations associated with the ownership of things both material and immaterial¹⁵. I believe that property plays a much more radical role within Rawls's theory than Cohen has anticipated.

First, Rawls notes that "the right to own certain kinds of property" is not a basic liberty and should not be protected as such¹⁶. For example, in relation to economic organization, Rawls states that "the means of production may or may not be privately owned," signaling an openness to a spectrum of economic ownership arrangements¹⁷. Unlike basic liberties, laws that dictate what kinds of property, "kinds" being differentiated by social function (e.g., economic production, housing, leisure), can or cannot be owned are not an unchanging constant within Rawls's framework. In a Rawlsian society, property ownership can be categorically created, altered, or abolished.

Beyond the right to own, Rawls has demonstrated a remarkable plasticity to the institution of property itself. While expounding on the concept of moral worth, Rawls states that "the institution of property ... is established for prior and independent social ends"¹⁸. He clearly conveys that the institution of property is organized to achieve social goals, goals which are morally prior to property itself. This not only entails the possibility of organizing property rights like the right to own or the right to income around social ends, but it also entails the malleability of property *obligations*, or the legal responsibilities that property owners have to other members of society.

The full thrust of Rawls's position is not to be underestimated. To address Cohen's argument, let us take up the institution of the family once more and explore how property law could help to reduce unjust inequalities within the home.

¹⁴ Cohen provides no definite solution to this dilemma, and rather points to the possibility of forming a social ethos that facilitates social justice yet that is not dictated by law or structural composition. See *Action*, pp. 24-30

¹⁵ I wish to provide a functional notion of property as opposed to a strict ontological theory.

¹⁶ *Theory*, p. 54

¹⁷ It is worth noting that Rawls still advocates for a market-based economic system, regardless of the ownership categorization employed by a Rawlsian society. See *Theory* p. 57

¹⁸ *Theory*, p. 275

The residential home has long been a space of oppression for women. The very maxim that the personal is political is a direct critique on the state's reluctance to interfere in what has traditionally been christened as "the man's castle." This property classification of the home as an exclusively "private residence" has made much of what happens within the home fall outside of political and economic purview. A clear example of this legally manufactured erasure is shown through our property relation to domestic labor. Since men are expected to perform the role of breadwinner within the family, mothers often forfeit their careers to raise children and perform domestic labor within the household. Because domestic labor occurs within the home and outside of the state's purview, both the state and the market have failed to recognize how essential domestic labor is to economic productivity.

Part of the problem lies in the legal understanding of private residences and the various social functions which they serve. In a bid to eliminate sexist inequalities, a Rawlsian society could reconceptualize its legal classification of family homes to also recognize their social function as sites of economic productivity. As long as domestic labor continues to be consigned to the status of a luxury good instead of a basic necessity, some women will continue to remain at home to perform unpaid labor in a bid to avoid the exorbitant costs of childcare and in-house cleaning. Recognizing residential homes as a sort of specialized extension of the economic market would entitle women who are performing domestic labor to receive private or public compensation.

More radically, a Rawlsian society could bring some forms of domestic labor completely outside of the home. Through mandatory zoning, the state could require government-sponsored childcare centers, dry-cleaning facilities, and other domestic labor establishments to be built in residential areas. Access to these sites could free both women and men from the domestic tasks that keep them from pursuing their careers. By partially reorganizing its system of property around domestic labor, a Rawlsian society could not only increase female participation in the workforce and reduce gender wealth gaps, but more fundamentally, they could enforce a public recognition that women are engaging, and have been engaging, in socially and economically productive work.

This public recognition can chip away at the problem that social coercion poses to distributive justice. By changing the public understanding of women and their role in society, we can create different expectations of what type of work women should be doing. The conceptual shift from "housewife" to "worker" better situates women in the political world outside of the private, non-political, and non-economic home. In this case, the conceptual shift was not caused by a social awareness campaign; it was caused by legal modifications to the property rights and obligations attached to housing and labor. These hypothetical modifications work to illustrate my central claim against Cohen's objection: discriminatory social expectations are not immune to the legal and material interventions prescribed by Rawls's theory.

To be fair, the solutions I have proposed do not completely resolve the problem highlighted by Cohen. Even with "business homes" or market remuneration for domestic labor, social coercive practices could still force women to disproportionately forfeit their professional careers. One could imagine a society with a guaranteed right to work, equal income, and near-perfect material equality that still suffers from sexist practices. Fundamentally, Cohen's arguments highlight how personal choice can throw a wrench into legislative approaches to distributive justice. But it is clear that Cohen has underestimated how dependent the execution of our behavior is on our legal structure. Through tactful legal maneuvers, legally coercive institutions can exert a significant corrective influence onto the social expectations and behaviors that cause unjust distributions.

Yet, from a Rawlsian perspective, one might object that wielding property law to affect personal behavior is just a round-about way of regulating basic liberties, violating the spirit behind Rawls's liberty principle. The goal of the liberty principle is to provide the most extensive scheme of civil liberties possible—it seems unlikely that such a malleable conception of liberty could be considered "extensive" in the sense required for the protection of liberal democracy. After all, if the state were to try to reduce gender discrimination by legislating away something as fundamental as the private residential home, then one could certainly argue that the functional consequence of such legislation is akin to violating one's right to privacy or some under fundamental liberty.

I am going to attempt to answer this objection by examining how the right to free movement operates within different transportation infrastructures. In a highway system, free movement entails that a driver going down the interstate can freely decide which exit she departs at, as well as the speed she travels and the lane she wishes to drive in. It is important to note that the exercise of her movement is dependent on the existing roads available to her. If we were to place that driver in a public transit system, one could hardly say that her right to free movement has been unjustly restricted. Rather, the exercise of her right has changed due to the new context we have placed her in. Instead of picking which lane to drive in, she must now decide which train to take, or which seat and cart to occupy. Her fundamental right to travel has not been infringed upon, it has only been reconceptualized.

Similarly, changes to our property relations do not necessarily infringe on our basic liberties. How we exercise of our rights depends on the material and social context that this exercise occurs in. While legal modifications to property will invariably impact the performance of our behavior, the claim that one's basic rights have been infringed upon is not equivocal with the claim that a previous behavior that used to be acceptable is not acceptable any longer. As evidenced through the right to free movement, a change in material context does not eliminate the ability to exercise a fundamental liberty, it only redefines how this exercise is to be accomplished in a new material environment. In this way, the fundamental liberties encompassed within the liberty principle can be preserved and protected as new social and material developments change the infrastructure in which they operate.

V. Social Coercion in the Original Position

Even if we grant the possibility that legal interventions into the basic structure can ameliorate socially coercive institutions, Rawls's theory might yet fail to provide a rationale by which a Rawlsian society should prescriptively choose to recognize and rectify social injustice. Social inequalities, as opposed to strictly economic ones, are harder to quantify and seem to be particularly invisible to a conception of distributive justice restricted to legal institutions. Despite this apparent incompatibility, I contend that Rawls's theory is particularly apt to address concerns of social inequality.

Rawls's conception of justice, which he calls "justice as fairness," considers "the social bases of self-respect" to be a "primary good" on par with other goods like wealth and liberty¹⁹. As such, to uncover how a Rawlsian society might identify social inequality, it is important to first understand Rawls's conception of justice and the mechanism he calls "the original position."²⁰

As a conception of justice, justice as fairness dictates that the structure of society should be agreed upon by rational and self-interested individuals on an equal standing. By forming an "original agreement" that will decide the arrangement of the society they will live in, these individuals act as their own representatives who will negotiate for the arrangement that will most benefit their interest. To ensure equality in negotiations, Rawls insists that representatives should have no knowledge of their future place within society or any of their defining characteristics by implementing what he calls "a veil of ignorance." Through the veil of ignorance, the representatives are to become unaware of economic and social features like class, ability, intelligence, race, gender, and so on, for both themselves and for others.

By forcing representatives to hold no particular social identity, Rawls argues that the veil of ignorance helps to secure the formation of a fair agreement in the arrangement of society's basic structure. Without knowledge of their social and economic standing, the representatives cannot exploit outstanding power imbalances or coercive factors to influence negotiations for their own benefit. In addition, because representatives have no set position within social arrangements, it is possible that, after the negotiation's are finalized and the veil is removed, the representatives could belong to any position in society, including that of the least-well off. Being self-interested, the possibility of inhabiting the most disadvantaged social and economic classes ensures that the representatives will attempt to maximize the interests of these classes in their negotiations. In sum, this hypothetical, pre-social position marked by equally-situated bargaining is what Rawls defines as the original position.

The original position has multiple functions within Rawls's theory. On one hand, Rawls uses the original position to derive his main principles of justice

¹⁹ See "Justice as Fairness," *Theory* pp. 3-40, p.54

²⁰ Rawls provides a detailed explanation of the original position in "The Original Position" *Theory* pp. 102-160

and to expand on his conception of justice as fairness. But Rawls also recommends a more practical application of the original position. Rawls imagines that one could inhabit the original position when undertaking deliberations to see whether a particular policy or structural arrangement is just. For example, if one were to consider the application of a steep regressive tax, one could inhabit the original position to see if a potentially low-income representative might rationally agree to the consequences of being taxed at a higher rate than a potentially higher-income representative. If the two equally situated representatives, while negotiating for their own self-interests, could both agree to such a policy, then the policy would be considered just under Rawls's conception of justice.

As the mechanism by which social inequalities are to be rectified, many, including Cohen, have found fault with Rawls's original position on the grounds that its principled abstraction from real social conditions makes it particularly incapable of handling identity-based social discrimination²¹. Due to the veil of ignorance, the existence of social identities like gender or race do not present themselves as intelligible considerations to the representatives in the original position. And if a social identity like gender cannot be considered within the original position, it seems improbable that gender-based discrimination can be recognized or addressed in negotiations between genderless representatives. Of course, taking Cohen's mode of abstraction, this oversight could surely apply to all forms of social identity and coercion, not just those which revolve around gender.

While Rawls's inattention to social identity has drawn him much ire, I do believe that Rawls's theory is often unfairly framed as being inattentive to social inequality. In actuality, Rawls's notion of distributive justice is highly sensitive to concerns of social equality, so much so that Rawls's inventory of "primary goods," the rights and advantages at the core of distributive justice, includes the social value of self-respect²². And not only does Rawls acknowledge the primary importance of equal social standing, but he also proposes certain legal interventions by which social self-

respect may be promoted within society. In discussing the assessment of education, Rawls writes:

The value of education should not be assessed solely in terms of economic efficiency and social welfare. Equally if not more important is the role of education in enabling a person to enjoy the culture of his society and to take part in its affairs, and in this way to provide for each individual a secure sense of his own worth²³.

Clearly, Rawls has a distinct appreciation for the aim of social equality, as well as for how certain policies affect social standing and engender social conceptions of self-worth. Rawls goes as far to say that social conceptions of self-worth might be "equally if not more important" than economic considerations, a valuation which gives significant insight into just how social inequalities are weighed within Rawls's theory.

Drawing from Rawls's discussion on education, we can confirm that it is possible to use the original position to rectify inequalities of primary goods within society. And if social equality and self-respect are to be considered primary goods, then it follows that social inequality, even if instituted by socially coercive institutions, would be of principal interest for the representatives within the original position. Although identity-based social inequalities are caused by identities which are invisible from the perspective of representatives in the original position, the inequalities (e.g., lack of self-worth or equal social standing) are themselves identifiable, paving the way for the recognition of social injustice under the veil of ignorance. While the liberty principle would bar representatives from directly regulating socially coercive behaviors that perpetuate social injustice, representatives could weigh and implement creative, non-liberty infringing legal interventions that would result in a redistribution of self-respect and social equality. Through this method, Rawls's inclusion of self-respect within primary goods provides a rationale by which a Rawlsian society should proactively identify and rectify social injustice caused by identity-based discrimination.

²¹ For a feminist critique of the original position, see Okin, Susan Moller. "Justice and Gender." *Philosophy & Public Affairs*, vol. 16, no. 1, Jan. 1987, pp. 42–72. Charles Mills has also advanced a racial critique of Rawls's race-blind approach to distributive justice in *Black Rights/White Wrongs: The Critique of Racial*

Liberalism. Transgressing Boundaries: Studies in Black Politics and Black Communities. Oxford University Press.

²² For a description and inventory of primary goods, see *Theory* p. 54

²³ *Theory* p. 87

VI. Conclusion

By exploring the role of property in Rawls's framework, I have shown that a Rawlsian society could use legal institutions to correct distributive injustice caused by sexist practices within the home. Beyond sexism, Cohen's critique is applicable to a broad range of discriminatory social conventions. Social expectations that shape personal behavior in light of race, ability, class, and other facets of one's identity are all of genuine concern for those who seek a comprehensive theory of distributive justice.

While I do believe that the problem of personal behavior remains a genuine difficulty in Rawls's theory, I also maintain that legal institutions, with creative and issue-specific lawmaking, can act to correct the unjust distributive consequences garnered by personal behavior. Beyond the capability of intervention, Rawls's theory also prescribes the proactive identification and elimination of unjust social inequalities through the classification of social self-respect as a primary good. Using the original position, a Rawlsian society could identify unjust social inequalities, leading to legal interventions aimed at a redistribution of primary social goods. For these reasons, I conclude that Rawls's basic structure can safely escape the predicament of Cohen's dilemma.²⁴

Works Cited

- G. A. Cohen. "Where the Action Is: On the Site of Distributive Justice." *Philosophy & Public Affairs*, vol. 26, no. 1, Jan. 1997
- Okin, Susan Moller. "Justice and Gender." *Philosophy & Public Affairs*, vol. 16, no. 1, Jan. 1987, pp. 42–72.
- Rawls, John. *A Theory of Justice*. 2nd ed., Belknap Press, 1999.
- Rawls, John. *Political Liberalism*. The John Dewey Essays in Philosophy: No. 4. Columbia University, 1993.

²⁴ I would like to thank the Prometheus Undergraduate Journal and the audience at the 2022 Mid-Atlantic Undergraduate Philosophy Conference for their comments and constructive suggestion.

Language and Common Sense in Berkeley's Philosophy

LONG NGUYEN

HILLSDALE COLLEGE

Abstract: This paper discovers a way to read two theses in the philosophy of George Berkeley, idealism and immaterialism, as commonsensical. This interpretation, as I will argue, is plausible because it fits well with Berkeley's view on language. Were the argument to be successful, it would vindicate, at least in part, Berkeley's insistence that he is a philosopher of common sense.

I. Introduction

A great deal of scholarship on the philosophy of George Berkeley is concerned with how it can be considered a common sense philosophy. In what follows, I characterize the philosophy of George Berkeley as the conjunction of idealism and immaterialism. Let *idealism* be the thesis that all that exists are either ideas or minds containing ideas, and *immaterialism* the thesis that physical matter doesn't exist. Now, despite Berkeley famously maintaining in defense of himself that he is a philosopher of common sense, "I side in all things with the Mob" (PC 405), it is not immediately clear how both theses can be commonsensical. As Seth Bordner (2017) points out, there is a dissonance among scholars over understanding the relationship between Berkeley's philosophy and common sense. Stances range all over the spectrum, from considering Berkeley's appeal to common sense simply as a rhetorical device (Yandell 1995), to rejecting Berkeley's philosophy as a commonsense philosophy (Bennett 2001), to accepting it as a defense of, or at least consistent with, common sense (Pappas 1999, Stoneham 2002), or to holding it as common sense itself (Luce 1967).

In this paper I make a case for understanding Berkeley's philosophy as commonsensical. I contend that in its commitment to common sense, idealism preserves the common use of language: it allows us to use language as we usually do. This stands in contrast to the usual reading of a common sense philosophy, which holds that such a philosophy is based on a set of intuitive, non-inferentially accepted propositions. It is under this later reading that Berkeley's philosophy

falls short as a candidate of common sense philosophy, as both idealism and immaterialism are inconsistent with this set of propositions. Distinguishing these two readings of common sense philosophy will thus be the key to understanding idealism as commonsensical. Immaterialism, on the other hand, can be understood as the result of Occam's Razor: the existence of the external world need not be posited to use language as we commonly do. What's more, I suggest that this reading of Berkeley is plausible because it fits well with Berkeley's view on language.¹ As a quick outline, I suggest that in its commitment to common sense, idealism adheres to our daily use of language.

In section 2, I identify what is usually meant by common sense philosophy, which should then shed light on the failure to recognize Berkeley's philosophy as commonsensical. In section 3, I present an alternative reading and provide textual evidence for interpreting Berkeley as such. In section 4, I investigate how idealism and immaterialism can be commonsensical under this reading. In section 5, I briefly put Berkeley into the greater historical context to see how this reading reveals interesting methodological insights between Berkeley and his contemporaries.

II. The Common Reading of Common Sense

By common sense philosophy, we usually mean common sense realism, the philosophy acquainted with Thomas Reid and later, G. E. Moore. For our purposes, it suffices to characterize common sense realism by its adherence to a set of propositions, each of which is noninferentially justified. Call these propositions *commonsensical propositions* (hereafter *C-propositions*). For example, propositions that physical, mind-independent objects exist and that we know of such objects directly through our senses are usually considered C-propositions.² By noninferentially justified, I mean a C-proposition is not justified by inference to another proposition yet is so well justified, such that it is really hard to overturn it, if possible. Consider two points for clarification. First, the skeptic may very well find accepting a proposition without justification naïve and unphilosophical. This ready acceptance is then liable to the accusation of dogmatism. In response, the common sense realist would say that to doubt a C-proposition is absurd and unintelligible: when one doubts a C-proposition, one ceases to be rational or reasonable in one's inquiry. "If anyone *was* so

¹ See also Bordner (forthcoming) and Pearce (2017).

² For a more detailed discussion on Berkeley and common sense realism, see Pappas (1991), Van Cleve (2017).

sceptical as to deny that any proposition is self-evident," writes Reid, "I don't see how we could convince him by reasoning" (EIP 6.4, original emphasis). Second, despite being taken as true unquestionably, a C-proposition is not a tautology; it expresses an empirical fact. That said, a philosophy is not a common sense philosophy when at least one of its theses is found inconsistent with this set of C-propositions.

Both immaterialism and idealism are inconsistent with the set of C-propositions. Following the list of C-propositions in Bordner (2011), they are inconsistent with the propositions that

C1 There are ordinary (macro) physical objects, and that

C2 All (macro) physical objects and at least some of their qualities exist independently of and are generally unaffected by perceptions.

While immaterialism is the negation of C1, idealism implies not-C1 because ordinary physical objects are neither ideas nor minds containing ideas. As C2 entails C1, both immaterialism and idealism are inconsistent with C2 also. C2 a fortiori conflicts with idealism: given idealism, all that exists is affected by perceptions.³

As we characterize Berkeley's philosophy as the conjunction of immaterialism and idealism, it is thus inconsistent with the set of C-propositions. Therefore, Berkeley's philosophy is not commonsensical. Straightforward as it may seem, this brief analysis reveals the central obstacle undermining Berkeley's philosophy as a common sense philosophy.⁴

III. The Alternative Reading: Common Sense in Practice

Common sense realism is not the only way to characterize a common sense philosophy. Another way to characterize common sense philosophy is by its adherence to *the common practices*. According to this characterization, a common sense philosophy preserves relevant norms—belief forming, knowledge acquiring, testifying, for instance—in play in the daily life. Putting it this way, the contrast to the usual reading is apparent: while the usual reading characterizes

common sense by its adherence to a set of *propositions*, this reading defines it in terms of *practices*. Also, note that here it's not my intention to argue for any rigid definition for the common sense, or common sense philosophy: I only suggest that this way of understanding common sense is intuitively plausible. Neither is it my intention to argue for or against any of those understandings common sense. My claim is that the alternative understanding accommodates another interpretation of Berkeley, which better justifies his insistence that he is a philosopher of common sense.

Berkeley, quite uncontroversially, can be read as preserving the common way of using language. Throughout his writings, Berkeley consistently condemns philosophers for confusing themselves by misconceiving the nature of language and is apologetic for the common, or "vulgar", way of speaking. In the Introduction of his *Treatise*, he sets out his goal to track down the source of errors in man's judgments and identifies such a source to be the wrong use we make of our faculties. To demonstrate this point, consider his famous dust-kicking analogy: "Upon the whole, I am inclined to think that the far greater part, if not all, of those difficulties which have hitherto amused philosophers, and blocked up the way to knowledge, are entirely owing to ourselves—that we have first raised a dust and then complain we cannot see" (PI 3). According to Berkeley, a specific misuse of faculties is the abuse of language (PI 6).

What does Berkeley mean by the abuse of language? Two relevant aspects of it are to undermine the capacity of language and to misconceive what is communicated. As he points out, "it is a received opinion that language has no other end but the communicating our ideas, and that every significant name stands for an idea" (PI 19). This is usually referred to as the *ideational theory of meaning*, hereafter the ideational theory. No sooner than introducing it, Berkeley makes it quite clear that he opposes this view. Notice that the ideational theory is presented as the conjunction of two theses: the one concerns the end of language, the other meaning. He rejects both.

Concerning the end of language, Berkeley thinks that the ideational theory impoverishes the capacity of language. He writes,

³ This conflict would be cashed out more clearly if one were to take Berkeley's famous doctrine *esse est percipi (aut percipere)*—to be is to be perceived (or to perceive)—instead. This doctrine is indeed very close to idealism: to be perceived is to be an idea, while to perceive is to be a mind.

⁴ A salient attempt to undercut this obstacle is due to George Pappas. As a quick overview, Pappas' proposal is that idealism is consistent with this reading of common sense, or at least to a greater extent than its materialism opponents, precisely because it is consistent with more C-propositions than these competitors are. See Pappas (2018), chapter 7.

Besides, the communicating of ideas marked by words is not the chief and only end of language, as is commonly supposed. There are other ends, as the raising of some passion, the exciting to or deterring from an action, the putting the mind in some particular disposition—to which the former is in many cases barely subservient, and sometimes entirely omitted, when these can be obtained without it, as I think does not unfrequently happen in the familiar use of language.

PI 20

Here, it can be observed that Berkeley gestures rather strongly toward a more robust view of language. While there are cases where language does communicate ideas, they only occupy the minority: we use language to achieve a lot more than mere communication. Concerning the meaning of what is communicated, or signs, Berkeley disagrees that ideas are communicated directly through signs. Given the diversity of language uses, the communicated ideas are *not* the same as signs taken literally. In fact, sometimes ideas are explicated from signs equivocally. To demonstrate this point, the following is worth quoting at length.

... In such things we ought to “think with the learned, and speak with the vulgar.” They who to demonstration are convinced of the truth of the Copernican system do nevertheless say “the sun rises,” “the sun sets,” or “comes to the meridian”; and if they affected a contrary style in common talk it would without doubt appear very ridiculous. [...] In the proper affairs of life, any phrases may be retained, so long as they excite in us proper sentiments, or dispositions to act in such a manner as is necessary for our well-being, how false soever they may be if taken in a strict and speculative sense. Nay, this is unavoidable, since, propriety being regulated by custom, language is suited to the received opinions, which are not always the truest. Hence it is impossible, even in the most rigid, philosophic reasonings, so far to alter the bent and genius of the tongue we speak as never to give a handle for cavaliers to pretend difficulties and inconsistencies.

P 51-52

The observation Berkeley points out here is that signs, when taken in an overstrict sense, result in misleading and even false ideas. However, they still successfully transmit appropriate ideas, precisely because we have conditioned ourselves as users of the language, such that we are not usually misled by taking signs literally.⁵

⁵ I thank Austin Heath for suggesting this point to me.

This point makes clear Berkeley's commitment to preserving common linguistic practices.

It can thus be argued that Berkeley's critique of the ideational theory gestures strongly toward the *use theory of meaning*, hereafter the use theory.⁶ According to the use theory, signs are meaningful not by virtue of certain ideas inherent in them, but rather by the way we *use* them in our linguistic practices. For instance, the meaning of “laptop” is not due to any inherent connection between the sign itself and smaller, tablet-like computers: its meaning is provided by our denoting such computers “laptop”. In other words, the idea associating with “laptop”, namely smaller, tablet-like computers, is determined by our use of the sign “laptop” itself. Putting it this way, the use theory is characterized by its divergence from the ideational theory in terms of its view on the relationship between signs and ideas. Consequently, its view on the end of language diverges also. As ideas are determined by our use of signs, language necessarily does more than just communicates ideas. This aligns almost perfectly with Berkeley's critique of the ideational theory.

IV. Idealism and Immaterialism Being Commonsensical

Before embarking on the task of understanding how immaterialism and idealism can be commonsensical, one has to recognize a central assumption in Berkeley's philosophical system, namely that all we perceive directly is ideas. Call this *the direct perception thesis*. He points out,

It is indeed an opinion strangely prevailing amongst men, that houses, mountains, rivers, and in a word all sensible objects, have an existence, natural or real, distinct from their being perceived by the understanding. But, with how great an assurance and acquiescence soever this principle may be entertained in the world, yet whoever shall find in his heart to call it in question may, if I mistake not, perceive it to involve a manifest contradiction. For, what are the aforementioned objects but the things we perceive by sense? [And] what do we perceive besides our own ideas or sensations?

P4

The direct perception thesis is plausible, at least in the historical context: Berkeley's supposed opponents, Descartes and Locke, agree on this thesis also.

⁶ For a discussion on interpreting Berkeley as committing to the use theory and his metaphysics, see Roberts (2017).

Granted that Berkeley is after a version of the use theory for his view on language and the direct perception thesis, we should be in a good position to see how idealism works well with common linguistic practices, which are rather robust under the use theory. Contra the ideational theory, the proponent of the use theory holds that we understand the meaning of a certain sign not by taking it literally, but by tracking its use in common linguistic practices. And after expounding this use, we get the idea encoded in that sign. That explains why, for instance, Berkeley thinks that one convinced by the truth of the Copernican system would still say, "the sun rises," "the sun sets," and so on. By "the sun rises," she doesn't mean the sun rises literally; she instead means movements of the earth and the sun are such-and-such. In another case, waving doesn't have any literal meaning as a sign: instead, by waving, one intends to greet another. The idealist has no problem explaining such cases, or exchanges of signs in general. In a generic exchange of signs, the sender will successfully transmit the idea if he understands the use of communicated signs. The receiver, likewise, will successfully render the idea if she understands the use of communicated signs. According to the direct perception thesis, she then perceives the idea directly. The takeaway of use theory is that what is grasped doesn't necessarily have propositional content. Of course, by "the sun rises," the communicated idea has propositional content, namely the proposition that the earth moves such-and-such and the sun moves so-and-so. On the other hand, by waving, the communicated idea doesn't have any propositional content. And since idealism can account for both cases, it is reasonable to say idealism works well under this view of language.

One can now see how immaterialism can be consistent with common linguistic practices. In particular, it is redundant to posit the existence of the external world while everything else works as fine. In the framework of idealism, we have seen that one can use language in the usual way only from the existence of minds and ideas. The existence of the external world doesn't add any good to the picture. It, in fact, brings a burden. Since the existence of the external world doesn't logically follow neither idealism nor the use theory nor the direct perception thesis, one needs to prove it prior to implementing it into the system. Put it that way, immaterialism can be understood as the

idealist employing Occam's Razor. Occam's Razor, or the principle of parsimony, is the principle that the simpler theory should be prioritized. As such, since the usual way to use language is preserved in the idealism framework, which doesn't include the existence of the external world, immaterialism should be preferred over materialism, according to Occam's Razor.⁷

V. Berkeley in Context

This interpretation of Berkeley reveals some interesting insights between Berkeley and his contemporaries. In what follows, I suggest Berkeley can be thought of as a middle way between two camps regarding external world skepticism and common sense: one of Descartes, Locke, Hume, and Kant, the other of Reid.

Considering external world skepticism, Berkeley joins the former in requiring a proof before accepting that the external world exists. As shown above, the existence of the external world is not included in the philosophical system, within which language operates robustly as in the daily life. According to Occam's Razor, one need not and should not add the existence of the external world into the basis. Reid, on the other hand, disagrees. He takes that the world exists independently from one's mind to be basic, thus requiring no further justification. At this point, joining the former side, Berkeley also holds that the direct perception thesis is basic.

Considering common sense, Berkeley joins Reid in accepting that skepticism is a non-starter and setting to preserve the common sense. While Descartes, Locke, Hume, and Kant are variably skeptical of the existence of the external world, Berkeley and Reid can be said to be anti-skeptics yet in the most opposing way. While Reid holds that the external world exists, Berkeley holds that the external world doesn't exist.

VI. Conclusion

In this paper, I discover a way to read idealism and immaterialism in the philosophy of Berkeley as commonsensical from the perspective of his view on language. These two theses are commonsensical not in the sense that they are consistent with commonsense propositions, but that they adhere to common practices, namely the common way to use language. Interpreting Berkeley as gesturing toward a version of the use theory of meaning, I proceed to show that

⁷ This is not the only upset of Berkeley over materialism. The standard interpretation reads Berkeley as rejecting materialism because it eventually leads to skepticism, which he disagrees.

idealism accounts for daily life communication, or exchanges of signs. Given that idealism accounts for the daily life communication, positing the existence of the external world is redundant as it doesn't add anything to the explanation. Immaterialism emerges as a result of Occam's Razor. This interpretation puts Berkeley into an intricate middle way between two camps: one consisting of Descartes, Locke, Hume, and Kant, the other of Reid.

Works Cited

Primary Sources

Berkeley, George. 2019. *Philosophical Commentaries by George Berkeley: Transcribed From the Manuscript and Edited with an Introduction and Index by George H. Thomas, Explanatory Notes by A. A. Luce*. New York and Oxford: Routledge. (Abbreviated "PC")

———. 2011. "A Treatise Concerning the Principles of Human Knowledge." In *Modern Philosophy*, edited by Forrest E. Baird, 309-50. Upper Saddle River: Prentice Hall. (Abbreviated "P")

———. 2011. "A Treatise Concerning the Principles of Human Knowledge." In *Modern Philosophy*, edited by Forrest E. Baird, 300-9. Upper Saddle River: Prentice Hall. (Abbreviated "PI")

Reid, Thomas. 2011 [1785]. *Essay on the Intellectual Powers of Man*. Cambridge: Cambridge University Press. (Abbreviated "EIP")

Secondary Sources

Bennett, Jonathan. 2001. *Learning from Six Philosophers, Volume 2: Locke, Berkeley, Hume*. New York: Oxford University Press.

Bordner, S. Seth. 2011. "Berkeley's "defense" of "common sense"." *Journal of the History of Philosophy* 49: 315-38.

———. 2017. "Immaterialism and Common Sense." In Brook and Belfrage (eds.), 343-54.

———. forthcoming. "Berkeley on Common Sense." In *The Oxford Handbook of Berkeley*. Oxford: Oxford University Press.

Brook, Richard, and Bertil Belfrage, eds. 2017. *The Bloomsbury Companion to Berkeley*. London and New York: Bloomsbury Academic.

Luce, A. A. 1967. "Sensible Ideas and Sensations." *Hermathena* 105: 74-83.

Pappas, George S. 1991. "Berkeley and Common Sense Realism." *History of Philosophy Quarterly* 8: 27-42.

———. 1999. "Berkeley and Skepticism." *Philosophical and Phenomenological Research* 59: 133-49.

———. 2018. *Berkeley's Thought*. Ithaca and London: Cornell University Press.

Pearce, Kenneth L. 2017. *Language and the Structure of Berkeley's World*. Oxford: Oxford University Press.

Roberts, John Russell. 2017. "Berkeley on the Philosophy of Language." In Brook and Belfrage (eds.), 421-34.

Stoneham, Tom. 2002. *Berkeley's World: An Examination of the Three Dialogues*. Oxford: Oxford University Press.

Van Cleve, James. 2017. "Reid's Opposition to Berkeley." In Brook and Belfrage (eds.), 299-313.

Yandell, David. 1995. "Berkeley on Common Sense and the Privacy of Ideas." *History of Philosophy Quarterly* 12: 411-23.

New Sights and Signs: Street art as a Response to Willful Hermeneutical Ignorance

SABAH RAINEY

CENTRE COLLEGE

Abstract: Activists increasingly employ street art as a tool to affect political change; as such, street art is becoming more and more prevalent in cities. This paper explores one way that street art can respond to injustices. In particular, I examine its effect on the epistemic injustice described by Gaile Pohlhaus, which is willful hermeneutical ignorance. While I argue for some ways art in general can respond to willful hermeneutical ignorance, street art in particular is well situated to accomplish this through its public quality.

Keywords: Willful hermeneutical ignorance, Gaile Pohlhaus, street art

Introduction

Street art is becoming more prominent in our cities, and it is often used to enhance public spaces. However, it is more than just decorations of public spaces. Street artists frequently employ it as a tool to comment on the status quo or to express their displeasure at current events. Guerrilla graffiti artists have made art out of potholes to draw attention to the state of disrepair. Tatyana Fazlalizadeh expresses the wrong of catcalling through her portraits of women who have been catcalled. During the COVID pandemic, street art expressing the need for unity emerged all over the world. In 2020 street artists depicted people killed by police officers in support of the Black Lives Matter protests against police brutality in the United States.

This paper explores one way in which street art responds to and helps correct injustices. In particular, I argue that through its public quality, which alters our daily spaces, some street art can communicate new concepts to audiences in a way that helps correct for willful hermeneutical ignorance. Willful hermeneutical ignorance being a kind of epistemic injustice put forward by Gaile Pohlhaus in which marginally situated knowers' epistemic resources are dismissed by dominantly situated knowers. This process occurs through the art evoking emotions in the audience such

that audiences have sympathetic emotional responses that help them interpret the illustrated scenario in the same way as the art demonstrates. The sympathetic response makes the audience more receptive to the epistemic resources presented in the work. To make this argument, I will first look at what street art is and what special features it has that make it well positioned to interact with audiences in this way. I will then discuss willful hermeneutical ignorance and how it is that dominantly situated knowers can dismiss epistemic resources of others. The analysis of the two will provide the framework for my discussion on how art generally, and street art in particular, corrects for willful hermeneutical ignorance. I will conclude with an analysis of the work of some muralists in Louisville who created those murals during the 2020 protests in the United States.

I. Street Art

I begin by defining the concept of street art; for this paper I will use Nicholas Riggle's definition of street art from the article "Street Art: The Transfiguration of the Commonplaces." An essential component of street art that I focus on is its public situatedness. To help explain the significance of situatedness for activists, I draw on Sondra Bacharach's view of street art as activist in spirit.

Nicholas Riggle defines an artwork as street art "if, and only if, its material use of the street is internal to its meaning."¹ In other words, the fact of being in the street situates street art, so as to make the street internal to the meaning and interpretation of the work. A painting on canvas may go from a gallery to a museum to the home of a collector without having its meaning affected. In those cases, the work's physical situation is external to its meaning as opposed to internal to it.

Riggle's use of the term "the street," however, must be qualified as he does not merely refer to the physical street. A cursory definition is any outdoor and accessible public space, which includes sidewalks, streets, parks, and plazas. In addition to the spatial street, Riggle asserts the street has a socio-cultural function. Its cultural function rests in the degree to which the street functions as a space that facilitates the communication and self-expression of individuals among others.² Because the socio-cultural function of

¹Nicholas Alden Riggle. "Street Art: The Transfiguration of the Commonplaces." *Journal of Aesthetics and Art Criticism* 68, no. 3 (August 2010): 246. <https://doi.org/10.1111/j.1540-6245.2010.01416.x>.

² Nicholas Alden Riggle. "Using the Street for Art: A Reply to Baldini." *The Journal of Aesthetics and Art Criticism* 74, no. 2 (April 2016): 192. <https://doi.org/10.1111/jaac.12280>.

the street is particularly important, when he says the use of the street is internal to the meaning of the street, it is the use of space that facilitates self-expression which street art makes use of in a way that is internal to its meaning.³

For a work of art to make material use of the street means the art uses the street as an artistic resource. Riggle distinguishes between two kinds of artistic resources: physical artistic resources and contextual resources.⁴ For a work of art to make material use of the street, it must incorporate physical elements of the street into the work itself.⁵ A muralist who paints on the sides of buildings uses that space of the street as a canvas, and thus a physical artistic resource. For a work of art to use the street as a contextual resource means the work is located and displayed in the street.⁶ The muralist decided to situate their art in the context of the street rather than in a gallery or a museum. Since Riggle emphasizes the centrality of the socio-cultural street, that is a component quality that the works make use of. They use a particular kind of material space, which is the socio-cultural street.

Riggle does not consider government funded public art or corporate advertising to be street art. Though it may be in the street, these sorts of works are institutionally protected and thus do not share street arts commitment to ephemerality. Secondly, they alter the spaces of viewing in ways that street art does not. Riggle asserts that they become spaces for looking at the art, thus art spaces. In that sense, the street is not truly internal to the meaning, rather they just happen to be displayed in a public location which is the street.

Sondra Bacharach, in the article “Street Art and Consent,” argues for an understanding of street art as a consensual, which leads her to assert that street art necessarily engages in activism as it physically alters social spaces without permission.⁷ While Riggle does not argue for the necessary aconsensuality of street art, and neither will I, the insight of the activist quality to street art is an important one that I will incorporate into Riggle’s view. Since Riggle conceives of the street as a space of communication and self-expression, street art’s use of the street means the works necessarily engage in this communicative process. Themselves

audibly or visually in the space. Even if the goal of street art is to enhance the street rather than to undermine or subvert it, altering social space can be seen as aiming to bring about social change, if only minimally through the one act.

To summarize, Riggle contends that street art’s use of the socio-cultural street is internal to its meaning. In interpreting street art, examining the formal qualities is insufficient, for that would be failing to account for the important artistic choices that contribute to the essence and meaning of the works. With this understanding, it becomes apparent that the situation of street art in its environment and the ways it alters its environment contribute to the meaning of the work as such. I will focus on this understanding of street art as I move forward to look at the relationship between the audience and the art.

II. Willful Hermeneutical Ignorance

In the article, “Finding your voice in the streets,” Sondra Bacharach argues that street art can address the two epistemic injustices Miranda Fricker describes in her book *Epistemic Injustice*.⁸ Gaile Pohlhaus, however, argues that Fricker overlooks other instances of epistemic injustice that fit into neither category; as such, she presents a third kind of epistemic injustice called willful hermeneutical ignorance. A testimonial injustice occurs when a hearer wrongly fails to afford the due credibility to the speaker. A hermeneutical injustice occurs when an experience is obscured from an individual’s understanding due to a gap in collective epistemic resources. Willful hermeneutical ignorance occurs when a hearers refuses to acknowledge the epistemic resources developed by marginally situated knowers. I assert that while Bacharach is correct that street art corrects for the epistemic injustices Fricker delineates, it can also correct for the third type that Pohlhaus discusses.

Pohlhaus’ definition of willful hermeneutical ignorance is founded on two principles of feminist epistemology: that we are interdependent and situated knowers.⁹ We rely on epistemic resources to analyze, evaluate, and understand our experiences.¹⁰ Language, concepts, and criteria are examples of epistemic

³ Riggle, “Using the Street for Art,” 192.

⁴ Riggle, “Street Art,” 245.

⁵ Riggle, “Street Art,” 245.

⁶ Riggle, “Street Art,” 245.

⁷ Sondra Bacharach, “Street Art and Consent,” *British Journal of Aesthetics* 55, no. 4 (October 2015): 481, <https://doi.org/10.1093/aesthj/ayv030>.

⁸ Miranda Fricker, *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford: Oxford University Press, 2011), 1.

⁹ Gaile Pohlhaus, “Relational Knowing and Epistemic Injustice: Toward a Theory of ‘Willful Hermeneutical Ignorance,’” *Hypatia* 27, no. 4 (Fall 2012): 717, <http://www.jstor.org/stable/23352291>.

¹⁰ Pohlhaus, “Relational Knowing and Epistemic Injustice,” 718.

resources; when we apply a concept to an experience we have had, we utilize an epistemic resource. However, as tools for communication, epistemic resources do not exist within an individual, but rather exist among a group, which means they develop within a group.¹¹ Because epistemic resources are developed within a group, marginalized people who maintain their community have the chance to develop resources outside of the sphere of dominant knowers.¹² In that we rely on collectively developed resources, we are interdependent knowers.

However, to be useful in understanding experiences, they need to be reflective of lived experiences. If epistemic resources fail to adequately interpret an experience, we can develop new epistemic resources to account for those experiences.¹³ For instance, women developed the concept of sexual harassment upon realizing they shared an experience for which there was not an adequate term. Because of the tension between lacking a concept and having an experience, they developed a new epistemic resource. More precisely, the fact that they had the opportunity to get into a group and discuss their experiences allowed for the development of the new resource.

The need to develop new epistemic resources based on our experiences is an example of how we are situated knowers. Based on our social position, we encounter experiences and challenges that those in different circumstances may not encounter. Because of that, our experiences we develop habits of¹⁴ Someone from an upper-class background, for instance, likely has different concerns than someone from a working-class background. A marginalized individual is situated differently, thus attentive to different things, compared to an individual who is part of the dominant group. As such, they have experiences dominantly situated knowers do not have, and if they share a community, they can develop epistemic resources together.¹⁵

Good epistemic resources will reflect those habits and help us understand the sorts of things we tend to experience and are concerned with.¹⁶ Dominantly situated knowers, however, frequently have a monopoly on what gets included in epistemic resources. In the case of women developing a concept of sexual harassment, the term was not around because men make up

the dominant class and did not have the experiences that made that term necessary to the same degree women did. That situation is why dominantly situated knowers are able to dismiss the epistemic resources of marginally situated knowers. In short, since the situation of dominantly situated knowers does not permit them to experience the experiences of marginally situated knowers, they can easily dismiss those experiences. Additionally, since the epistemic resources of marginally situated knowers are not immediately useful to dominantly situated knowers, it can be difficult for them to see the value in them or to fully incorporate them into their lives.¹⁷

I ought to briefly discuss the wrong of willful hermeneutical ignorance. Willful hermeneutical ignorance disadvantages marginalized groups when dominantly situated knowers refuse to acknowledge their resources. Pohlhaus identifies a few specific harms of willful hermeneutical ignorance. One being that “it blocks the transmission of knowledge that ought to make a normative claim on those for whom the knowledge is intended.”¹⁸ That being the case, there is a sense in which willful hermeneutical ignorance prevents the communication of ideas that could contribute to the political discourse. It is the fact of being willfully ignorant that makes it so that people cannot engage with the resources.

III. Street Art as a Response to Willful Hermeneutical Ignorance

In this section I argue that art can respond to willful hermeneutical ignorance; the public situatedness of street art enables it to responded effectively to willful hermeneutical ignorance. I begin by discussing how works of art generally can give viewers the chance to employ some of the artist’s epistemic resources. I will then discuss the qualities of street art that make it particularly suited for accomplishing this task.

Art and Communication

Works of art evoke emotional responses in audience members rendering them more receptive to and sympathetic towards the viewpoints expressed in the art. When viewers share in an artist’s interpretation of depicted events, they make use of the artist’s epistemic resources. Through experiencing new viewpoints,

¹¹ Pohlhaus, 718.

¹² Pohlhaus, 731.

¹³ Pohlhaus, 718.

¹⁴ Pohlhaus, 717.

¹⁵ Pohlhaus, 719.

¹⁶ Pohlhaus, 719.

¹⁷ Pohlhaus, 722.

¹⁸ Pohlhaus, 732.

viewers have the opportunity to experience the utility of concepts in interpreting certain experiences.

In order to argue street art can correct for willful hermeneutical ignorance in some measure, I need to be clarify what counts as correcting for willful hermeneutical ignorance. It is too high of a bar to require that street art solve willful hermeneutical ignorance in order to count as an effective response. When Bacharach wrote her article arguing for street art's ability to respond to the testimonial and hermeneutical injustice, she did not argue that street art had to entirely solve those injustices in order for it to still be an effective response. Instead, she argues that epistemic justices "would involve correcting for these negative identity prejudices, enabling the marginalized voices to speak and for their voices to be acknowledged and heard".¹⁹ The key criteria for her is that street art provides a platform for marginalized knowers to express their knowledge in such a way that it is validated, not that street art solves all instances of injustice. Thus similarly, I do not need to argue that street art can and does solve all instances of willful hermeneutical ignorance. Since willful hermeneutical ignorance is primarily the case of voices being unacknowledged, the most relevant part of Bacharach's criteria to this type of injustice is the criteria that it allows for the voices of marginalized knowers to be acknowledged. However, it is since willful hermeneutical ignorance can and does affect the ability of knowledge to enter the public discourse, it is also important that the voices are heard, that they give marginalized knowers the chance to contribute knowledge to the public discourse. To tie that more specifically to willful hermeneutical ignorance, it means artists that are marginalized knowers have the chance to share their epistemic resources with a wider audience and that those resources are acknowledged and heard.

Artworks often evoke emotions in us. In doing so, it can make us more receptive to what the artist is illustrating. Young argues that artworks can evoke memories of personal experiences in audience member; in doing so, artworks can get the audience to remember particular details that then can help support the viewpoints expressed in the art.²⁰ Artworks can evoke those kinds of memories through emotional evocation as well. When in the same or a similar

emotional state as the work, we might remember related experiences and become more receptive to the hypotheses of the work, we become more receptive to the artists perspectives. In a sense, we reflect back at the work what the work has reflected onto us. When we engage in this process of reflecting back onto the work, we allow ourselves to interpret the imagery in the way the artist intended. If a scene is depicted as tragic, we pick up on that and through feeling sad ourselves are more open to the idea that the scene is in fact sad. When we are taking on a similar emotional state to the artist, or towards the art, we are taking on an interpretive lens. Through that process, we make use of the artist interpretation of an experience. When we see the painting as portraying a sad state of affairs, we are interpreting the events depicted as sad. To make that sort of evaluative judgement is to make use of an epistemic resource as epistemic resources are tools for evaluating and understanding experiences.

If street art must share epistemic resources in order for it to be effective, I need to discuss what it means to share resources and how art can do that without actual words. We interpret events and objects based on our experiences with and understandings of those objects. Therefore, if we are to interpret art the way the artist intended, we have to have some shared understanding of the signs and symbols being employed by the artist.²¹ The gold circle around someone's head is interpreted as a halo for us, but we could imagine another culture that interpreted it radically differently. This is to say that for an image to communicate ideas and understandings the artists intend there has to be some shared understanding of symbolism. When the artist BMike paints a mural of a girl with a halo around her head (fig. 1) and her arms opened and turned upwards in a receptive gesture, he incorporates imagery associated with saints to elevate her and present her as sacred. Were audiences not aware of that symbolism, the image would not be understood. When street art uses shared symbolism to communicate concepts or

¹⁹ Sondra Bacharach, "Finding Your Voice in the Streets: Street Art and Epistemic Injustice," *The Monist* 101, no. 1 (January 2018): 35, <https://doi.org/10.1093/monist/onx033>.

²⁰ James Young, "Literary Fiction and True Beliefs," *Art and Belief*, ed. Ema Sullivan-Bissett, Helen Bradley, and Paul Noordhof (Oxford: Oxford University Press, 2017), 96.

²¹ Paul C. Taylor, *Black Is Beautiful : A Philosophy of Black Aesthetics* (Hoboken : Wiley, 2016), 44.



Figure 1 Odums, Brandon "BMike." Paint. Royal St., New Orleans, LA.

evaluative judgments, it communicates epistemic resources. If these resources are outside those commonly employed by dominantly situated knowers, there must still be some overlapping frame of reference, otherwise the attempt at communicating would fall flat. In the case of epistemic resources, this means that since street art is communicating different epistemic resources to an audience that may be unfamiliar with such resources, they have to make sure to communicate in understandable ways. To tie it into what I was saying earlier, in order to evoke the appropriate emotional response in the audience that would allow for the effective communication of meaning, the artist has to use signs/symbols that are shared and understood by both parties to evoke those types of emotional responses. This process corrects for willful hermeneutical ignorance in two ways. It gets resources into the public domain, thus allowing for the spread of resources to a large audience. Additionally, the epistemic resources are temporarily being used, which may allow people to, at least momentarily, perceive them as useful. If part of what allows someone to dismiss new resources is that they do not have the necessary experiences to understand the relevance of the new resources, then giving those people the opportunity to experience how the new resources are used by and pertinent to people can help them see the value in these resources.

I have argued that art can provide viewers with the chance to make use of new concepts and tools for evaluation, epistemic resources, and in doing so, art can help viewers see these resources as valuable. This is accomplished through evoking an emotional state in the viewer that makes them more receptive to views and concepts expressed in the work. In doing so, art

works correct can correct for willful hermeneutical ignorance because they put resources into the public domain and because they help get those resources acknowledged and valued by others.

Why Street Art?

The key feature of street art that well situates it to correct for willful hermeneutical ignorance is its public quality. That means that it reaches a broader audience and literally aims to change the public environment. In reaching a broader audience, street art makes the epistemic resources it presents to more people, which gets the resources heard by more people. In changing the public environment, street art makes the epistemic resources it presents a constant feature of the public spaces where it acts in.

Street art is placed in the public domain in a significant way; after all, it is put in public spaces such that the audience does not have to ask to see it or even get the choice to see it. This differs significantly from art in museums or galleries; while both of those locations are accessible to the public, perhaps even free to the public depending on the circumstances, they are not as available. Street art reaches more people and a wider array of people. It is placed in a public space and is seen by more than just the people who visit museums. Because of those factors, it reaches a broader audience than other forms of art. This is an advantage because it places the epistemic resources that the artist conveys in their art to a wider group of people. Since part of the harm of willful hermeneutical ignorance is linked to the fact that it prevents the resources from reaching other people, having the art placed in a public space that is readily accessible to lots of people where they mean to access it or not helps correct that harm of willful hermeneutical ignorance.

An additional advantage of street art for artists is that street art circumvents institutional barriers that may exist. This circumvention is significant in two ways: there isn't control of the artists' work or of how it is portrayed and more artists have the chance to get involved with street art. That is, with street art there is no, or very little, curatorial oversight, input and/or control. Because of this, the artist has a level of freedom to decide how and in what context their work is displayed. This gives them more artistic autonomy and freedom, which is an advantage when displaying marginalized epistemic resources. The second way street art circumvents institutional barriers is by allowing more artists to access the medium of street art. That is, since street art does not need curatorial approval,

anyone can engage in the practice regardless of education or background in the subject. Now, while it may be the case that it takes practice and knowledge to become skilled at street art, it does not require an arts education, or for a gallery to sponsor and approve of someone. As such, we might say there are fewer barriers to becoming engaged in street art practices.

An additional factor I would like to consider that street art does is that it is public in quality because of the fact that it aims to alter the public environment. Let us break this down. Street art engages with the public domain significantly—through being in a space that is a public space of communication and self-expression—as such it physically alters the public sphere. This quality is something that Sondra Bacharach defines as being “activist in spirit.” Let us analyze what precisely she means by this. While she does not quite define the terminology of activist, we can take the word to mean something like aiming to enact social or political change. As such, when Bacharach argues that street art is activist in nature, we might take it to mean that since street art literally changes the social space, it is necessarily activist to some degree. While we might debate whether in fact all street art is activist in this way, it seems that we can comfortably say that street art aiming to promote certain world views or perspectives certainly would fit this goal of being activist in nature – it changes the physical, social environment in a way that also aims to present certain new perspectives to the public.

I would additionally like to return to the significance of the fact that street art constantly exposes people to the new epistemic resources and changes the public and social environment of places. Faegin discusses the significance of artworks in shaping the environments in which we act. That is, paintings can “function to reinforce certain habitual modes of acting and thinking, rather than imparting knowledge.”²² Art creates certain spaces that foster certain ways of acting and thinking. For example, in the case of religious art, the art, when placed in the intended environment, makes present the divinity and evokes ritual behaviors.²³ A key feature here is that art is in its intended environment, not in a museum, at least in the case of art that was not intended for that situation. Street art is similarly positioned. That is, it was intended to be

looked at in the public context of the street, and it is not really removed from that context.

Now that we have considered how street art can be activist in spirit, let us consider the significance of that. In this case, part of being activist in spirit is physically altering the public environment. Street art alters the everyday spaces; it constantly bombards people with the art. That is, people are frequently and repeatedly exposed to the work of art, rather than going into a museum, looking at it and then moving on. This gives people the chance to be repeatedly exposed to the epistemic resources of the artists. I would argue this is an important way in which street art is activist in nature. That is, a quality of being an activist is working towards change without others necessarily wanting you to or asking you to. This relates well to street art. In museum art, viewers get the choice of whether or not to engage with the art, curators get the choice of whether or not to display the art. With street art, those types of choices are diminished if not entirely removed. Exposing people to something without really asking, simply acting is an important consequence of the argument that street art is activist in nature.

Street art, being activist in spirit as I have said, physically alters the public space. Because of that, it physically alters the space we act and live in. Therefore, part of being activist in spirit is that street art creates an environment meant to promote certain ways of thinking and or acting. It is meant to create a space that promotes certain epistemic resources, that promotes certain values. Since it makes that space when it physically alters the public environment. This differs from museum art which is removed from the daily lives of people and from the environments in which they frequently act.

In conclusion, the intensely public quality of street art, which is a way in which the art engages with public dialogue, enables the art to push its epistemic resources on more people, and it physically alters the space people live in. Altering the space makes it so that the epistemic resources of the work are constantly presented to audiences and not something that can be simply or easily overlooked.

IV. Case Study: Black Lives Matter Murals in Louisville, Kentucky

²² Faegin, Susan. “Paintings and their Places.” In *Art and Its Messages: Meaning, Morality, and Society*, edited by Stephen J. Davies, University Park, Pa.: Pennsylvania State University Press, 1997: 25.

²³ Faegin, “Paintings and their Places,” 22.

During 2020, there were nationwide protests across the United States in response to racial inequality and particularly in response to police brutality against African Americans. The protests were largely sparked after the death of George Floyd at the hands of police officers. In Louisville, murals sprang up all over the city, particularly after the death of Breonna Taylor, which occurred in March of 2020 in Louisville; many of the murals depict people who were killed by the police. Murals depicting the value of those who have been shot by the police demonstrate the belief of the artist that those lives are valuable and worth being displayed publicly in an overwhelmingly large manner. The portraits make the people larger than life and take a stance on what and who is valuable to have a portrait of and to be in our space. Again, street art makes a value judgement on how our space ought to be and what is worth being in the space. I will examine two murals in particular: the *No Justice No Peace Louisville* and a mural of Lady Liberty by Wilfred E. Sieg III.



Figure 2 Holbourn, Whitney, et al. *No Justice No Peace Louisville*. Summer 2020. Paint. 11th and Main St., Louisville, KY.

No Justice No Peace Louisville (fig. 2) by Whitney Holbourn, Andrew Thompson and Braylyn “Resko” Steward, located at the intersection of 11th and Main Street in Louisville, Kentucky depicts the faces of Breonna Taylor, George Floyd and others who have been killed at the hands of police officers. The mural depicts all of the figures as larger than life, literally; all of the faces are huge. This depiction makes these people who were killed in acts that deemed them as less valuable, as hugely valuable. That is, the mural makes these people the dominant features of the space. When someone drives or walks by, they cannot help but look at the mural if only because of the size and the colors of the work. It stands out and brings these people into the public space. To do so demonstrates the views of the artist that these people have value and are worth portraying. These are portraits in some sense, and to make

a large portrait that is put into public space inherently demonstrates the opinion that they are of value. This contrasts with the way that they were treated at the hands of police officers.

All of the features are highlights in a variety of colors, and because of the colors the mural itself is eye-catching. These two features inspire feelings of wonder and awe in the audience. Once again, such feelings contribute to the fact that the mural dominates the space. But also, in a sense, to have feelings of awe is to have respect and a degree of value for what is portrayed. When the audience is made to feel awe, they are primed to see these figures as people of value. In doing so, they adopt the value judgement of the artist, which contributes to accepting and making use of the artist’s viewpoint and resources.

The second mural I will discuss is the mural by Wilfred E. Sieg III (fig. 3). This mural makes use of more complicated visual symbolism to transmit the idea that African American lives are treated with less value and that this state of affairs is contrary to American ideals. The artist uses the American flag and the statue of liberty as representations of American values and ideals. These are well understood symbols in the United States and for many they may evoke feelings of patriotism and national pride. However, the artist portrays Lady Liberty in a state of despair. Her torch is lowered, and her hand covers her face as if she were crying or feeling ashamed. Through depicting Lady Liberty in such a state, the artist evokes similar feelings in the audience: we understand the state she is in as a reflection of that which the country is in and we may feel despair and sadness as well. Additionally, the artist associates Lady Liberty with the rights of African



Figure 3 Sieg, Wilfred E. III. Summer 2020. Spray paint. 543 S. Shelby St., Louisville, KY.

Americans through covering her in black power symbols. Her torch makes up the stand for the scales in the art. The scales of justice are another symbol in the United States that we understand as a symbol of the fairness of the justice system. In other words, they are another symbol of the way in which the United States is fair and just. The scales, however, are out of balance. The gun, which is still smoking, is heavier than the body of the African American. The implication being that in our country we do not value the lives of African American people. Guns and those who use them are more important and worthwhile than those who are killed by them.

Through the subversion of typical symbols of patriotism, liberty, and equality the artist plays on feelings of national pride to provoke audiences into reflective and shamed states. The national symbols are portrayed incorrectly as it were and in such a way that should make us uncomfortable, it should cause us to feel that something is wrong. Since this use of symbolism evokes that kind of emotional response in the audience, the audience is primed to see the situation displayed as wrong. If the audience has a similar sympathetic response, they are in a better position to accept the proposed hypothesis. In this case the view is that there is in fact a state of affairs where the lives of African Americans are not valued and that this is a bad state of affairs which we should seek to change.

I have presented two examples of street art in Louisville, Kentucky that effectively convey viewpoints of the artists. In doing so, these works give audiences the chance to make use of the epistemic resources presented. The artists use works of art as a tool for activism and protest of the status quo. Through putting their art into the public sphere, they not only seek to reach a wide audience, but they alter the space and demonstrate what kinds of values they would like the social environment to reflect.

Conclusion

I have argued that street art can be used by artists to correct for the particular epistemic injustice of willful hermeneutical ignorance by giving audiences the opportunity to experience the epistemic resources of the artists. Street art does this through evoking a sympathetic emotional response in viewers, which makes the viewer more receptive to the proposals the art illustrates. Through giving audiences the chance to experience how the artist interprets certain scenarios, the audience gets the opportunity to make use of epistemic resources they may not otherwise have

experienced. This Helps correct for some of the harm of willful hermeneutical ignorance.

While it is certainly the case that any good art may have a similar effect, that is, it can give its audiences the chance to make use of and experience new concepts, street art is particularly well suited for this task. This is not to say that all street art does this. Some street art may respond to entirely different sorts of injustices, or it may be largely apolitical with the simple goal of beautifying the social environment. However, street art's public position gives it a large viewership, and it alters the space we act and live in so as to constantly make those resources present to the audience. As such, street art has special qualities that make it well-positioned to correct for willful hermeneutical ignorance.

Works Cited

- Bacharach, Sondra. "Finding Your Voice in the Streets: Street Art and Epistemic Injustice." *The Monist* 101, no. 1 (January 2018): 31–43. <https://doi.org/10.1093/monist/onx033>.
- . "Street Art and Consent." *British Journal of Aesthetics* 55, no. 4 (2015): 481–95. <https://doi.org/10.1093/aesthj/ayv030>.
- Baldini, Andrea. "Street Art: A Reply to Riggle." *The Journal of Aesthetics and Art Criticism* 74, no. 2 (2016): 187–91. <https://doi.org/10.1111/jaac.12261>.
- Faegin, Susan. "Paintings and their Places." In *Art and Its Messages: Meaning, Morality, and Society*, edited by Stephen J. Davies, 17–25. University Park, Pa.: Pennsylvania State University Press, 1997.
- Fricker, Miranda. *Epistemic Injustice : Power and the Ethics of Knowing*. Repred. Oxford: Oxford University Press, 2011.
- Holbourn, Whitney, et al. *No Justice No Peace Louisville*. Summer 2020. Paint. 11th and Main St., Louisville, KY. *Ride Spot*, <https://ridespot.org/rides/186150>.
- "Louisville artists unveil mural honoring Breonna Taylor; others killed by police." *Wave 3 News*, Updated 6 July 2020, <https://www.wave3.com/2020/07/06/louisville-artists-unveil-mural-honoring-breonna-taylor-others-killed-by-police/>. Accessed 11 December 2021.
- Odums, Brandan "BMike." Paint. Royal St., New Orleans, LA. *The Helis Foundation*, <https://thehelisfoundation.org/wp-content/uploads/2020/04/5d138eae5a7a7.image.jpg>
- Pohlhaus, Gaile. "Relational Knowing and Epistemic Injustice: Toward a Theory of 'Willful Hermeneutical Ignorance.'" *Hypatia* 27, no. 4 (2012): 715–35. <http://www.jstor.org/stable/23352291>.
- Riggle Nicholas Alden. "Street Art: The Transfiguration of the Commonplaces." *Journal of Aesthetics and Art Criticism* 68, no. 3 (August 2010): 243–57. <https://doi.org/10.1111/j.1540-6245.2010.01416.x>.
- . "Using the Street for Art: A Reply to Baldini." *The Journal of Aesthetics and Art Criticism* 74, no. 2 (April 2016): 191–95. <https://doi.org/10.1111/jaac.12280>.
- Sieg, Wilfred E. III. Summer 2020. Spray paint. 543 S. Shelby St., Louisville, KY. *Courier Journal*, <https://www.gannett-cdn.com/presto/2020/06/11/PLOU/0498d818-c3f7-4d1a-a6ed-b881dab4a5be-BreonnaMural.JPG?width=660&height=452&fit=crop&format=pjpg&auto=webp>
- Taylor, Paul C. *Black Is Beautiful: A Philosophy of Black Aesthetics*. Foundations of the Philosophy of the Arts, 6. Hoboken: Wiley, 2016.
- Young, James. "Literary Fiction and True Beliefs." In *Art and Belief*, edited by Ema Sllivan-Bissett, Helen Bradley, and Paul Noordhof, 85–99. Oxford: Oxford University Press, 2017.

Reasoning Together

TOBY RIVAS

GRINNELL COLLEGE

Abstract: In this paper, I criticize the traditional liberal model of political discourse on the grounds that its individualistic conception of reason alienates political subjects from themselves and from their communities, and it can only justify that alienation by appealing to a formal relativism where all arguments are equally valid prior to rational evaluation. I then draw on Hannah Arendt's work to construct and defend an alternative model of political discourse, grounding it in an attentiveness to humans as fundamentally social creatures. This model introduces criteria for legitimate political reasons, letting us avoid the formal relativism and alienation of the liberal model without reducing politics to mere arbitrary preference or might-makes-right. It is a model of political discourse oriented towards human flourishing.

Introduction

In July 2020, *Harper's Magazine* published "A Letter on Justice and Open Debate," wherein more than 150 noted academic signatories voice their opposition to "a new set of moral attitudes and political commitments that tend to weaken our norms of open debate and toleration of differences in favor of ideological conformity."¹ This letter is one example of an increasingly popular critique of the political left: that they are opposed to rationality, tolerance, and debate, relying instead on social and institutional pressure to position their views as matters of moral fact that need not be defended by reasoned argumentation.

The letter's more affirmative claim is that "the way to defeat bad ideas is by exposure, argument, and persuasion," and to deviate from that norm is to sap the lifeblood of our liberal society.² This sentiment, and the letter in general, is significant because it reflects the standard liberal model of political discourse and action: humans are fundamentally rational, and they use their reason to appraise the world and form opinions. Those opinions then clash in the public sphere, rational debate leads one opinion to rise above the rest, and that

opinion compels legitimate political action. This framework uses the universal accessibility of reason to provide a collectively agreeable alternative to simple might-makes-right politics. As Jürgen Habermas puts it, rational argumentation is the "court of appeals" that resolves disagreements that cannot be resolved through everyday routines, that those disagreements not devolve into the exercise of direct or strategic force.³ Non-discursive responses to reasoned argument, then, are modeled as problematic because they seek to circumvent the rationality that grounds any sort of political action as legitimate. If someone says something I disagree with, and I tell them to be quiet (and my response is seen as legitimate in the public sphere), I have reduced the political to arbitrary preference. And so, it makes perfect sense that the letter insists on "the value of robust and even caustic counter-speech from all quarters."⁴

Yet this model, taken at face value, asks us to bite two problematic bullets: it asks us to disregard perlocutionary effects—the ways a given speech act impacts its listeners, distinct from the process of rational appraisal—when evaluating and responding to political speech, and it asks us to live in a political system where marginalized or historically marginalized groups can never feel fully secure in their equality or franchise. Suppose that I am attending a Black Lives Matter protest. Someone comes up to me and tries to engage me in argument, saying that Black people commit more crimes per capita than white people, so it is reasonable that they are shot by police. Under the liberal model, I have an obligation to debate them. By attending a protest, I am making a claim in the political sphere, and they have presented a reasoned (albeit poorly reasoned) argument against me. If I believe that I am right and they are wrong, I should present reasons to that effect. Indeed, under the liberal model, we "judge the rationality of a speaking and acting subject by how he behaves as a participant in argumentation," where a rational subject either concedes to the force of opposing reasons or replies to them, and an irrational subject is "deaf to argument."⁵ And here are the bullets. First, debating the racist likely carries a negative perlocutionary effect in that I am being asked to expend a great deal of mental and emotional energy to debate the very humanity of myself or my peers. It should be settled fact; I should not have to debate it.

¹ (Ackerman et al. 2020)

² (Ackerman et al. 2020).

³ (2004, 18).

⁴ (Ackerman et al. 2020).

⁵ (Habermas 2004, 18).

Second, by engaging in debate, I am acknowledging the fact that the political climate and system emerges merely from the sum of (purportedly rational) opinions held by the populace. If I want to ensure that Black people gain better treatment or maintain their existing rights, I had better argue for it. This relativism, necessitated by the liberal model, ensures that egalitarianism can never be secured within a political system, because it can always, *in principle*, be debated away by some new, sufficiently persuasive argument.

That is, the liberal model is *alienating*: it alienates us from ourselves insofar as we are social and emotional creatures, and it alienates us from our communities insofar as we seek a sense of belonging and mutual regard. The model asks us to disregard perlocutionary effects in our responses to political arguments because they are politically irrelevant. But to discard perlocutionary effects as politically irrelevant is to discard subjective experience and expression in general as politically irrelevant. We may of course draw on personal experience to shape and evince our opinions and arguments, but subjectivity as such is politically irrelevant. When we enter the political sphere, we are meant to become impartial, rational evaluators. If our subjective well-being conflicts with our stance as impartial, rational evaluators, we must either disregard that subjective well-being or temporarily remove ourselves from the political sphere. This is alienating. Similarly, the model's insistence that franchise and equality can always be debated away alienates us from our communities. We cannot feel truly secure in the mutual regard that binds a community together if that regard can be displaced at any moment through purportedly rational argumentation. Undocumented migrants cannot feel secure in the US, for example, because ICE's enforcement priorities swing wildly with the presidency, i.e., with the process of political debate. Any given group can, in principle, always be ejected from the community on the grounds of a sufficiently persuasive argument.

A proponent of the liberal model could very easily refute these two points of criticism. The perlocutionary effects argument falls easily before a counterexample. Suppose I am a white mother attending an anti-school integration protest. Much the same argument follows: it is obvious that integration will worsen the lives of my children; I am scared of Black people; I should not have to debate something that should be settled fact. Telling the mother that this subjectivity is politically irrelevant alienates her. The difference, of course, is that segregation should not be "settled fact." But if we

treat the perlocutionary objection to debate as legitimate, there is no way to know. We must engage in rational inquiry *before* we can know if perlocutionary objections to debate are legitimate, lest we use the objection as an excuse to entrench the status quo. It is the protestor's job to regulate their own emotional state, and that state should not interfere with legitimate political discourse. The second argument falls similarly: while progress and franchise can, in principle, be debated away, so can any form of oppression or injustice. And, because rational discourse tends towards good outcomes, we can rely on the free exchange of information and ideas to carry us along a liberatory social trajectory. If there is an oppressive ruling class, they may find it alienating that their social position can be debated away by rational argumentation, just as minority groups may find it alienating that their franchise can be debated away. This alienation is not a reason to reject free, rational argumentation.

Both responses rely on a relativism that stems from the liberal account's individualistic model of rationality. Recall that, under the liberal model, political discourse is the product of interplay between everyone's personal opinion. Everyone is equally rational and equally capable of insight, so we must formally treat every opinion as equally worthy of debate, since the only way to know if it is correct or not is to rationally investigate it. We may hear an argument, see it through, and then dismiss future instances of the argument (ignoring anti-vaxxers after gaining an understanding of the science, for example), but nothing in the form or content of any given opinion can cause us to hold it above or below any other opinion prior to substantive, rational evaluation.

This individualistic conception of political reason is the basic flaw of the liberal model. Rather than placing the rational, self-interested individual at the center of our politics, I propose that politics should attend to humans as Arendtian beings-with others, morally significant agents that are bound together in relations of mutual care and responsibility. Such a conception lets us model political discourse in a way that a) satisfyingly solves the liberal model's alienation problems, and b), lets us criticize bigoted arguments on their face without reducing politics to arbitrary preference or might-makes-right. I do not aim to criticize reason; I aim to reconcile our reason with our relatedness in the political sphere. The remainder of this paper will develop and defend that model.

I. Living Together

For Hannah Arendt, the most basic and significant condition of human existence is plurality, the fact that “men, not Man,” populate the earth.⁶ This condition is “the condition” of political life and thought, for politics is necessitated and motivated by the fact of human *interaction* (Arendt 2018, 7). If only one human existed, the political sphere would not and could not exist. Political systems are our responses to the condition of plurality, and political arguments are arguments about the efficacy or goodness of those responses.

Broadly, there are two routes we may take when confronted with human plurality: we may attempt to mitigate plurality, constructing political systems that allow everyone to go about their business, as unfettered as possible by others, or we may embrace plurality, constructing political systems that conceive of society as more than an aggregate of distinct individual interests. We will refer to the first approach as living in parallel (or a “parallel approach”), and to the second as living together.

The key feature of parallel approaches—of classical liberalism, for example—is that they are grounded in a methodological and ethical regard for the individual as such. Thomas Hobbes’s political thought is the archetypical example. He begins by imagining self-interested, rational, resource-consuming individuals in the state of nature, living side-by-side without any laws or established society. Here, Hobbes thinks people live in a state of perpetual war because the only way to protect oneself from violence in the state of nature is to have more resources than one’s neighbors, and so everyone must constantly compete to amass more and more resources. Hobbes then derives his political system by arguing that, in the state of nature, sufficiently free and rational individuals would enter into a social contract, because it is the only way to rise out of perpetual war and meaningfully secure their interests.⁷ Society and the state, then, develop and derive legitimacy from the interplay between self-interested individuals as such.

In other words, this approach insists that we can properly and adequately respond to the human condition of plurality without actually referencing humans as plural. Arendt calls this tendency the “avoidance of the public,” the tradition’s continual “attempt to run away specifically from the fact that *we are*, in the

plural sense, and are therefore responsible for each other.”⁸

The alternative approach, seeking to live together, views humans as essentially beings-with others, following Arendt. When humans are born, they enter into a relationship of dependence on their parents and community—dependence not just for biological needs but for social interaction, recognition, and development. Indeed, chronic neglect in children is associated with worse psychological damage than chronic abuse.⁹ And even in adults, loneliness is strongly associated with depression and suicidal ideation.¹⁰ Moreover, this human dependence on and regard for others is reflected in prehistory, the closest thing to a true “state of nature” that has ever existed. Rather than competing as self-interested individuals, even the earliest, least-developed humans cooperated with one another, demonstrated investment in social bonds, and even expended resources to care for the sick and disabled.¹¹

Where parallel approaches ground themselves in the individual as such, together approaches ground themselves in a regard for humans as beings-with others, deriving a system of responsibility from our essential relatedness. As Arendt puts it, by entering into a community, we are “charged with the care of others,” and others are charged with our care.¹² Rather than political arrangements existing to facilitate the parallel, self-interested operation of rational individuals, we may view political systems as oriented towards the flourishing of humans as beings-with others.

This discussion is relevant because it can help us better understand the traditional model of political reasoning. It is a parallel model, grounding itself in the individual as such. Politics, for the traditional account, is about the interplay of rational, self-interested individuals, so political reasoning is about interplay between individually held, rational opinions. Just as a political system and set of institutions emerge from the interplay of discrepant individual interests, legitimate, mutually amenable political action is thought to emerge from the interplay of individual opinions and beliefs. Two things here point us towards a solution. First, obviously, we want to develop a model of political discourse that is grounded in a regard for humans as beings-with others. Second, and more subtly, we can gain leverage by thinking about specifically political

⁶ (2018, 7).

⁷ (2008).

⁸ Arendt quoted in Schulman 2006, 80; Schulman 2006, 80. We will examine this responsibility in more detail momentarily.

⁹ (Harvard University, n.d.)

¹⁰ (Stravinsky and Boyer 2001)

¹¹ (Gorman 2012).

¹² (qtd. in Schulman 2006, 78).

reasons. The traditional model takes for granted that valid political reasons are nothing more than rationally justifiable opinions about political matters. Since the individual as such is at the center of politics, the individual's opinion is naturally at the center of political discourse. But if we place plurality at the center of politics, we have space to develop a concept of specifically political reasons, i.e., reasons that pertain to plurality.

Before continuing, note that I do not purport to prove, in the strict sense, that we should prefer living together to living in parallel. I cannot, in this paper, solve ethics by satisfactorily grounding a moral principle without relying on normative assumptions. Rather, I have given some reasons as to why I think we should live together—that the human experience is fundamentally and inescapably shaped by our relations with others and our embeddedness in communities—and as this paper unfolds, I think we will see that thinking in terms of living together is beneficial for our political models, especially in that it allows us to conceive of rational political discourse without alienating people from themselves or their communities. Ultimately, though, I am relying on a moral assumption. One may reject that assumption and subsequently reject this paper's arguments.

II. Political Reasons

If a reason is to impel action, it must reference some external motivating force. It must be grounded, somehow. Suppose it is raining, and someone tells me to put on a raincoat. I ask them why, and they say, "because it is raining." The rain is a reason to put on a raincoat *if and only if* I plan on going outside and do not like getting wet. This is a version of the is-ought gap. Reasons can contain facts about the world, and they can make logical inferences based on those facts, but no combination of facts and inferences can compel someone to act without appealing to some normative assumption.¹³ People can derive this normativity from a host of different sources: their personal religious or moral beliefs, their arbitrary personal tastes, their valuing of their own lives or happiness, and so on. Different people value different things when making

decisions and evaluating reasons. Recall that, under the traditional model of political discourse, we must treat any given political opinion as legitimate, and we can only discard someone's political claim after we have defeated it in rational argument. By this token, the liberal model requires us to treat any given normative assumption as politically legitimate until proven otherwise. This is not to say that the traditional model does not care about how a given reason is grounded. Rather, because the traditional model grounds itself in and derives political legitimacy from the individual, normative positions held by individuals are *the* politically intelligible sources of normativity.¹⁴

This grounding scheme also implies a particular type of answerability. By engaging in political reasoning, under the traditional model, we make ourselves answerable to others insofar as they are rational agents. Reason must always leave itself open to debate and criticism, and since no one is omniscient, it must always leave itself open to criticism from others, insofar as that criticism is rational. We will return to the idea of answerability in a moment.¹⁵

If we want an account of political reasoning that sets plurality at the heart of politics, we must begin by figuring out how to ground our political reasons. We need a politically intelligible source of normativity. We cannot just say "plurality" or "our Arendtian responsibility to care for others" and leave it at that. Rather, we must remember that politics is about us attempting to live together, so properly political reasons are reasons that comment on how we can best live together. The actual "we" here is significant: "we" may be humanity as a whole, it may be us as Americans, or as students attending the same college, or nearly any other group. By affirming our membership in a particular we, we affirm a level of responsibility and regard for other members of the we. Political reasons, then, are reasons that draw their normative force from the regard and responsibility associated with membership in a particular we.

Here, we should clarify the concept of a we and the associated responsibility. I have chosen to use we as a noun, rather than choosing some synonym of "group," because it emphasizes the affirmation of membership

¹³ This is an oversimplification (see Hilary Putnam, "The Collapse of the Fact/Value Dichotomy"), but it is adequate in the context of our discussion.

¹⁴ Of course, some liberal thinkers ground their politics in some idea of rights or some universally acceptable standard. But even if they think some universal standard exists, insofar as they are liberal, they must understand all political claims as equally valid on

their face, prior to rational evaluation. This principle is most clearly articulated by John Stuart Mill in *On Liberty*, but I mean to get at the broad understanding of political discourse that rises out of the liberal tradition, exemplified by the *Harper's* letter.

¹⁵ This argument is inspired by a similar discussion of Laden's (2012, 14–15).

and responsibility. As we already established, all humans are unavoidably born into communities, dependent on others. When I say “we,” rather than saying “me and these people,” I am signaling my commitment to live together with them, affirming my responsibility to care for others and their responsibility to care for me, rather than running from the public and wishing to live as an individual in parallel with other individuals. Since the model we are developing sees living together as the goal of politics, we’s are our relevant political groups. We can picture Venn diagrams within Venn diagrams: the largest circle is all of humanity, and within it there are circles representing countries, local communities, and so on, where each circle is a particular group aiming to live together in a particular context or geographic region.

This in mind, we can articulate the characteristics of a valid political reason, i.e., the characteristics of a reason properly grounded in a we. First, as I already said, a valid political reason draws its normative force from the regard and responsibility associated with the we. Second, a valid political reason is answerable to other members of the we insofar as they are rational agents *and insofar as they are members of the we*. If someone says, “mass incarceration is good because Black people are morally bad,” their argument hinges on their willingness to ignore any Black person that responds, “actually, I am an equal member of society.” The argument requires an implicit reconstitution of the we in a way that excludes Black people, so that it can make claims that ignore Black people. This invalidates it as a political argument because political reasons pertain to a *particular* we, attending to its makeup and respecting its members’ commitment to living together. By ignoring the actual we, which includes Black people with whom we are committed to living together, it no longer properly pertains or attends to the we in question. By this token, we are not obligated to respond to bigoted ideas with reasoned debate, as they exclude themselves from justified political relevancy by refusing to ground themselves in a way that gives them bearing on the relevant political community.

Here, we should pause and note that I am making a normative claim about what *ought to count* as a political reason or argument. Of course, “Black people are worse than white people” is political in the colloquial sense, and it certainly pertains to politics. I am introducing criteria to distinguish between legitimate and illegitimate political claims. When I say a bigoted argument is not political, I mean that it is not *properly* political. It is not legitimate; it should not enjoy the

regard and consideration we afford to legitimate political claims.

We can further clarify our picture of political reasoning by distinguishing between 1st, 2nd, and 0th order political arguments. 1st order political arguments are about the constitution of the we: who is included, how it is arranged, and so on. 2nd order political arguments are about what goes on in the we. These are policy questions. The 0th order political question is whether there should be a we at all, that is, whether we should confront plurality by living together or by living in parallel. For a productive argument to be had at any given level, there must be agreement at the prior levels. If you and I are to debate policy (2nd order), we should agree that policy should be informed by a regard and responsibility for our peers (0th order) and we should agree on who the relevant peers are (1st order). Using this framework, we may say that racist policy positions stem from 1st or 0th order disagreements. They are not mere differences in opinion, they are either differences in attentiveness to the relevant we or differences in commitment to living together in the first place.

This model still has several problems. We lack a clear picture of what healthy discourse actually looks like within this framework, especially 1st-order discourse about the constitution of a given we. Moreover, there is nothing to stop racists from making sound 1st-order arguments in favor of exclusion, nor is there a picture of what constitutes a sound 1st-order argument in general. Finally, we should consider whether it is actually good to have a model that lets us exclude bigoted arguments from political discourse on their face. After all, if we want to live together, we should talk to each other.

III. Reasoning Together

We can begin to address these problems with our model by returning to the two problems of alienation we initially ascribed to the liberal model, beginning with its inattentiveness to perlocutionary effects and the associated alienation from our subjectivity. This is a problem of answerability: the traditional model sees arguments as answerable to others only in their capacity as rational agents who can present counterarguments, excluding them in their capacity as sufferers of adverse perlocutionary effects. But under our model, properly political arguments are also answerable to others in their capacity as members of the we, as morally significant agents for whom we are partly responsible. Moreover, we are politically answerable *to ourselves* as morally significant agents.

Therefore, a political argument can be criticized on the grounds that it produces serious, adverse perlocutionary effects in its listeners, and anyone can reasonably refuse to engage in debate by way of politically legitimate attentiveness to their own wellbeing. This is not to say that anyone can shut down a political argument by complaining that it hurts their feelings. Rather, we must remember that the goal of political argument is to better live with our peers, so political arguments that harm those peers in their presentation or content are counterproductive. Discourse may be attentive to this fact without being stifled.

That is, when we make political arguments, we should be actively affirming our commitment to living together with those around us. By attending to our arguments' perlocutionary effects, the act of political argumentation is itself an act that constitutes and strengthens the we. If we instead ignore perlocutionary effects, or even take pride in our ability to offend our political opponents (see, for example, conservatives selling paraphernalia branded with "liberal tears"), the very act of political argumentation damages the we by rejecting the relationship of mutual care that ought to ground political discourse.¹⁶

The broader point, though, is that this way of thinking undoes the traditional model's alienation. Rather than drawing a strict and adversarial distinction between people as emotional, experiential subjects and people as rational political actors, we align the subjective and universal within the political realm by grounding the political in a regard for people as morally significant, socially embedded bearers of responsibility—agents defined by their subjective landscape just as much as by their rationality. Politics is about people, and people are more than their rationality. By embracing that fact, our model of political discourse avoids alienation and better helps us live with one another as people, rather than as merely rational, self-interested individuals.

Our model also solves the second alienation problem, stemming from the fact that franchise and equality are always subject to being debated away by a sufficiently persuasive argument. As already discussed, our model solves this problem by letting us discard bigoted arguments as politically invalid by virtue of their inattentiveness to the relevant we. The difference between this and the traditional approach is that the traditional approach models political discourse merely as a way to mediate between everyone's individual opinion, where

any of those opinions may come out on top. Our model gives political discourse a telos, and that telos is incompatible with arguments for whatever subjugation, provided the potentially oppressed party is already part of the we. Thus, we bind equality and franchise to the political community itself. Debating away equality and franchise would require shifting a society's entire political paradigm, moving from one aimed at living together to one aimed at living in parallel. This is of course possible, especially given some catalyzing event or a large enough timescale. But by conceptualizing arguments for oppression as politically invalid, rather than as merely poorly-thought-out opinions, we can create a scheme of political discourse that does not alienate the vulnerable from their peers or from politics itself.

Above, I said that the potentially oppressed group must already be part of the political community. This leads us to one of the remaining problems with our model: we lack an account of how we's are to properly expand or contract, and of what valid 1st-order arguments (arguments pertaining to the makeup or arrangement of a particular we) look like in general. Since the goal of politics is to live and flourish together, the goal of any particular we is to facilitate that broader goal. Arguments about the constitution of a we, then, should pertain to its efficacy or appropriateness as a group of people committed to living together within a particular context. An argument about efficacy, for example, could be made in favor of limited governmental jurisdiction, because it is logistically difficult to properly govern huge numbers of people across vast swaths of land. An argument about context might pertain to a community's size, purpose, geographic location, and so on. For example, Grinnell College routinely rejects applicants because it is a small, academically rigorous college with limited housing, space, and faculty. To function properly, it must be able to exclude certain people from membership.

Why, then, can I not (for example) construct an ethnostate and declare that we must keep non-whites out to function properly? The key is that we's should rise out of the broad commitment to living together, as opposed to living in parallel. When Grinnell rejects an applicant, it is not rejecting them as a human being. Indeed, Grinnell emphasizes its philosophical and institutional commitment to living together with people all over the world, not just with those in its immediate

¹⁶ (Prestigiacomo 2018).

academic community. Rejection from an ethnostate, though, signals a total rejection, a denial of any commitment to living together in general, in favor of a belief that we can strictly define human societies, stratify them based on moral worth, and live in parallel. Thus, a valid, properly political 1st-order argument rises out of a general commitment to living together and aims to illuminate how the we in question may better facilitate that goal.¹⁷ An ethnostate is necessarily a state predicated on living in parallel, not together.

Here, one might wonder how we are to understand, say, a serial killer, who we might imagine as warranting exclusion from the we. No action of an individual, no matter how evil, can warrant their wholesale exclusion from the we. If someone murders someone else, it is reasonable to incarcerate them, or at least implement some sort of coercive intervention. That intervention will likely involve a reduction in their liberty, a lowering of their social status, and so on. But even murders maintain their humanity, and we must maintain the basic commitment to living together with them. Under my model, a claim like “we should torture serial killers because they have given us license to exclude them from moral consideration” is politically invalid, and that is a strength of the model. There may be legitimate arguments for the jailing, disenfranchising, and general differential treatment of serial killers. But if we aim to live together, no one, through any heinous action, can exclude themselves totally from the basic commitment to living together.¹⁸

None of this is to say that we should never speak to racists, or sexists, or people who lack a commitment to living together. If we want to live together it is crucial that we talk to one another and that we leave room for people to develop their outlooks over time. We can understand this by drawing a distinction between political discourse and general moral or social discussion. It can be perfectly appropriate, even productive and morally venerable, to discuss someone’s racist views with them in a personal capacity. To borrow from

Anthony Simon Laden, reasoning between individuals, properly conceived, is a process of mutual attunement and discussion aimed at reaching a shared understanding.¹⁹

It is healthy and important to engage in this sort of talk about one’s moral convictions and political beliefs. Rather than being focused on proselytization or on achieving a particular political goal, it provides a space for people to engage with their peers as worthy partners in moral thought, developing their outlook and way of thinking. Racially prejudicial arguments are not politically valid and should not be seriously considered in the political sphere, but it is entirely reasonable to talk with someone who harbors racial prejudice.

The key difference between personal discussion and political discourse is that in political discourse, participants conceive of themselves as making political claims that, if accepted, should have implications on the political community’s behavior or constitution, whereas in personal discussion, participants understand themselves as trying to better understand one another. They may discuss things that have political implications and may make arguments that would be invalid if presented as political arguments, but they are not aiming to make an argument to the political community or otherwise directly impel political action by virtue of their argument. A participant may hope that she persuades her interlocutor and that that persuasion has a political impact, but the political impact does not stem from the broader political community being involved or persuaded. The goal is not to win an argument or impel action but to further mutual understanding.

Conclusion

We have thus found a way to give our understanding of political discourse meaningful moral direction without falling into repression or dogmatism, and we have done it by understanding people as morally significant bearers and subjects of responsibility, as beings-with

¹⁷ Again, note that I do not purport to *prove* that living together is the morally correct approach. I would not expect this paragraph (or paper) to persuade a white nationalist, and it is not meant to. Rather, I mean to provide a more useful way of understanding political discourse that is more morally congruent with my (and I hope with my readers’) ethical perspective than the traditional model.

¹⁸ Here, my opinion differs significantly from Hannah Arendt’s. In *Eichmann in Jerusalem: A Report on the Banality of Evil*, Arendt says that Eichmann must hang because his violent assertion of the authority to reject the Jewish people as partners in sharing the Earth renders any commitment to living together with him,

sharing the earth, unreasonable and perhaps self-defeating. Eichmann’s actions, I think, warrant his exclusion *up to* the point of the violent excommunication from humanity that Arendt advocates. We may strip his citizenship and condemn him to live the remainder of his life in an international prison. But, if we aim to live together, we are not entitled to carve out an exception to that aim for one class of person, nor are we entitled to declare that someone is no longer worthy of humanness. I feel no sadness that Eichmann hung, but Arendt’s philosophical justification is lacking.

¹⁹ (2012, 8).

others. This is an achievement because morality, orientation towards human flourishing, is *good*. The traditional model alienates us from ourselves and from our communities because it asks us to be answerable only to rationality, not to ourselves or to each other as social, morally significant beings. We avoid this alienation and gain a generally useful descriptive and prescriptive model when we build our politics on a commitment to living together.

Of course, I do not purport to have solved this issue over the course of a 6000-word paper. There are still many ways to defend the liberal model against my critique, and one could certainly levy a variety of objections against my model, particularly in its reliance on Arendt's account of responsibility and its arguable inattentiveness to the everyday realities of modern political discourse. Routes for further investigation remain.

But we have constructed a descriptively and prescriptively useful philosophical lens. The model is descriptively useful because we can better understand sentiments like those in the *Harper's Magazine* letter, seeing non-discursive responses to bigoted political arguments not as threats to the free and open exchange of ideas, but as rejections of politically invalid arguments that deny our commitment to live together. And this model is prescriptively useful because it helps us understand and strive for a politics that is genuinely attentive to human plurality. We thus begin to understand how to reason together if we wish to live together.

Works Cited

- Ackerman, Elliot, Saladin Ambar, Martin Amis, Anne Applebaum, Marie Arana, Margaret Atwood, John Banville, et al. "A Letter on Justice and Open Debate." *Harper's Magazine*, July 7, 2020. <https://harpers.org/a-letter-on-justice-and-open-debate/>.
- Arendt, Hannah. 1964. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Viking Press.
- . 2018. *The Human Condition*. Chicago: University of Chicago Press.
- Gorman, James. "Ancient Bones That Tell a Story of Compassion." *New York Times*, December 17, 2012. <https://www.nytimes.com/2012/12/18/science/ancient-bones-that-tell-a-story-of-compassion.html>.
- Habermas, Jürgen. 2004. *Reason and the Realization of Society*. Vol. 1 of *The Theory of Communicative Action*, translated by Thomas McCarthy. Cambridge, MA: Polity Press.
- Harvard University. n.d. "Neglect." Center on the Developing Child. Accessed April 22, 2022. <https://developingchild.harvard.edu/science/deep-dives/neglect/>.
- Hobbes, Thomas. 2008. *Leviathan*. Edited by J. C. A. Gaskin. Oxford: Oxford University Press.
- Laden, Anthony Simon. 2012. *Reasoning: A Social Picture*. Oxford: Oxford University Press.
- Mill, John Stuart. 2011. *On Liberty*. Cambridge: Cambridge University Press.
- Prestigiacomio, Amanda. "It's Here! 'The Daily Wire' Officially Launches Merch Store." *Daily Wire*, July 23, 2018. <https://www.dailymag.com/news/its-here-daily-wire-launches-merch-store-amanda-prestigiacomio>.
- Schulman, Stephen. 2006. "Yes, She Is an Ethicist: Arendt, Responsibility, and Existentialism." In *Existentialist Thinkers and Ethics*, edited by Christine Daigle, 70–90. Montreal, QC: McGill-Queen's University Press.
- Stravynski, Ariel, and Richard Boyer. 2001. "Loneliness in Relation to Suicide Ideation and Parasuicide: A Population-Wide Study." *Suicide and Life-Threatening Behavior* 31, no. 1 (Spring): 32–40. <https://doi.org/10.1521/suli.31.1.32.21312>.

Plato's Tripartite Classification of Normative Reasons

WEIMING SHENG

VASSAR COLLEGE

Abstract:

"I am still unable, as the Delphic inscription orders, to know myself; and it really seems to me ridiculous to look into other things before I have understood that",¹ says Socrates in *Phaedrus* (230a). The idea of self manifestly fascinated Plato, and this paper discusses one of his most well-known ideas: the division of the soul into reason, spirit, and appetite in the *Republic*. Overall, I focus more on the division itself than on the result of the division. §1 introduces two main lines of interpretations for Plato's tripartition. §2-5 show why the division does not establish multiple subjects in the soul, and §6-8 give reasons for why the theory cannot plausibly be about motivating reasons. §9 gives the conclusion that the tripartition should be taken as a classification of normative reasons.

I. Two Interpretations of the Tripartition

Plato's tripartite division of the soul² into reason (τὸ λόγιστον), spirit (τὸ θυμοειδές) and appetite (τὸ ἐπιθυμητικόν) in Book IV of the *Republic*³ could either be understood as a theory about the soul, or as a theory about human action. The two general directions have overlaps, and commentators often hold them both.⁴ I think Plato's tripartition is about human action, and in this essay, I argue for the plausibility of the view that Plato's division offers a tripartition of normative reasons.⁵ I expound my view by undermining two other popular and plausible alternatives: in §2-5 I argue

against the claim that Plato successfully uses what is by convention called *the principle of opposites* to establish multiples subjects or agents within the soul. If this is right, Plato's tripartition cannot be taken as a theory about the constitution of the embodied soul based on his arguments. In §6-8 I argue that Plato's theory does not work very well if it is taken to be about motivating reasons, both in the case of mental conflicts and in general. I give my conclusion and briefly assess the results of this paper in §9.

II. The Three-Parted Soul

First, let's say the tripartite division is supposed to be a claim about the constitution of the soul. More specifically, it would be the claim, prevalent in the scholarship, that Plato makes the soul out to be an amalgamation of three agencies or subjects, each with its own agency-granting structure,⁶ and each with desires, cognitive capacities, and pleasures attributable to it.⁷ No matter which part bears precisely which attributes, this reading takes Plato to claim, in a non-metaphorical sense, that the human soul is constructed out of these three agent-like parts. There are two reasons for reading Plato's argument in this way.

One reason is that the language used in the relevant passages strongly suggests this claim. Introducing "the difficult part" of the inquiry, Socrates⁸ asks whether "we do each thing *with the same thing* or, since there are three, we do one *with one* and another *with another*" (436a-b).⁹ Here, "with the same thing", "with one", and "with another" translate "τῷ αὐτῷ" and "ἄλλῳ". These datives in Greek are most plausibly read as datives of instrument,¹⁰ and a case can be made that what Plato puts in the dative of instrument for an act implies its being the subject of that act.¹¹ This

¹ From Alexander Nehamas and Paul Woodruff's translation.

² The ancient Greeks' idea of the soul (ψυχή) is free of much of the conceptual baggage that it carries for us. Julia Annas suggests that "Plato's talk of 'soul' is rather like our talk of someone's 'mind' to refer to certain mental phenomena without committing ourselves thereby to any particular theory about the mind" (124). I accept this: hence, in this essay, when I speak of the "soul", I mean, broadly, "that part of a man's personal life with which he identifies himself" (Crombie 34).

³ Another division occurs in Book X, based on what Jessica Moss calls "cognitive dissonance" and "emotional conflict" (35). I won't discuss that here because I think the psychological theories in Book IV and Book X are different. There are also other divisions in the corpus: notably, in the *Phaedo* and the *Timaeus*. I do not discuss those here because I think the version Plato puts forth in the *Republic* is sufficiently mature and sophisticated, and it is intelligible by itself.

⁴ Hendrik Lorenz, for example, thinks that Plato's theory is at once "about the nature of the embodied human soul and a theory of human motivation" (*Plato* 243).

⁵ That is, Plato offers three groupings, categories, or headers under which at least some of the normative reasons for our actions can be put.

⁶ The structure guarantees that each part is not simply a collection of desires. See Irwin 217-220.

⁷ See Bobonich 219-220; Lorenz, *Brute* 25.

⁸ Socrates is the main character in most of Plato's dialogues. To what extent, and in which dialogues, do his views represent Plato's are under dispute. I assume in this paper that in the *Republic* he is used as a mouthpiece for Plato's own views, and thus I sometimes use "Socrates" and "Plato" interchangeably.

⁹ Passages of the *Republic* are from C. D. C. Reeve's translation with some variations.

¹⁰ van Emde Boas et al., 378.

¹¹ Burnyeat, 29-39.

seems to suggest that Plato is dividing the soul into subjects or agents. More directly, when discussing what justice “in truth” is, Socrates claims that the just person “harmonizes the three parts of himself”, “binds together those parts” and “becomes one from having been many things” (443b). This seems to show that the soul is divided literally into parts. If Plato, then, did not think that the tripartition should be applied to the constitution of the soul, delineating multiple subjects or agents within it, why would he have expressed the division in these terms?

Another reason for favoring this reading is that for Plato's defense of justice to work on the level of the individual, the division must be in a literal sense about the constitution of the soul. Jon Moline puts this forcefully: unless Plato admits to having wasted three and a half books “exploring a blind alley”,

nothing less than a structural and functional *isomorphism* between polis and psyche will do...The psyche must upon examination prove to be literally complex, and literally reducible to parts which are independent of one another in the sense that they can stand in direct conflict as political factions do.¹²

That is, since Plato has divided the city into three in a non-metaphorical way, his division of the soul must also be a claim about its constitution in a non-metaphorical way, if there is to be an isomorphism between them. But there must be an isomorphism between them for his defense of justice to work. Therefore, Plato's claim is about the constitution of the soul.

These two reasons are external in the sense that they are reasons why *Plato*, writing the dialogue, might have taken the division as one about the constitution of the soul; they are not internal to the argument itself, and thus they are not good reasons for *us*, the readers, to take the division in this way. To see whether we should accept that the soul is three-parted, it is necessary to examine the arguments which Plato gives for the division.

III. The Principle of Opposites

The basis of Plato's division is the principle of opposites: “the same thing cannot do or undergo opposite

things; not, at any rate, in the same respect, in relation to the same thing, at the same time.” (436b). The principle thus claims that, under the conjunction of the same subject (ταὐτόν), same respect (κατὰ ταυτόν), in relation to the same thing (πρὸς ταὐτόν), *and* the same time (ἄμα), opposites cannot be done or undergone.¹³ The principle itself, however, does not specify how it can be decided, when opposites are done or undergone, *which* condition or conditions in the conjunction are to be negated.¹⁴ Therefore, without as already given which conditions are *not* to be negated, even in the case of opposites, none of the conditions can be validly negated based on the principle itself.

Thus, if, as Richard Robinson suggests,¹⁵ “the skeleton of the argument” for the division is the principle of opposites and the fact that sometimes people are thirsty but not willing to drink, nothing can be said as a result about whether there are two distinct subjects in the soul, one wishing to drink and another forbidding it, without having first established that the opposites occurred in the same respect, in relation to the same thing, and at the same time. But it is impossible for Plato to establish the condition of “in relation to the same thing” in this case, if he also wants to posit what Robinson calls the *principle of relatives*: “whenever things are related to something, those that are of a particular sort are related to a particular sort of thing, whereas those that are just themselves are related only to a thing that is just itself” (438a-b). Thus, for example, thirst itself, unqualified, is related to drink itself, and not to something qualified such as a tasty drink or a healthy drink (439a-b). Or, in more abstract terms, if set X stands in a certain relation with set Y, it does not follow that X stands in that relation with any proper subset Z of Y. But if we posit the principle of relatives, it would not then be in relation to the same thing that the person both wants and does not want to drink, but in relation to drink itself that they want to drink, and in relation to something else about the drink, say its poisonousness, that they do not want to drink. Since the conditions are given in conjunction, as long as one of them is negated, the conclusion that the occurrence of the opposites distinguishes two subjects cannot be

¹² Moline, 2-3.

¹³ The prepositions “κατὰ” and “πρὸς” can have more or less the same sense when followed by an accusative, which is the case here, but obviously there must be a distinction since they are mentioned here as two separate conditions—though in fact some editors have been tempted by this to delete “πρὸς ταὐτόν”! However, as James Adam explains, “κατὰ ταυτόν is ‘in the same part of it’...while πρὸς ταυτόν is ‘relatively to the same thing, viz. to something

other than the subject of the proposition” (246-247). This agrees with Reeve's translation and with the second apparent counterexample Socrates responds to at 436d-e, which requires a distinction in “κατά”.

¹⁴ This point is made by A. W. Price (41), albeit, I think, somewhat obscurely.

¹⁵ Robinson, 44.

validly affirmed. Since one condition, “relation to the same thing”, is in fact negated, opposites thus *can* occur without positing two subjects.

The problem exploited by this counterargument is inherent in the simultaneous postulations of the principle of opposites *and* the principle of relatives, *not* in any examples that Plato happens to use to illustrate his point. As long as an unqualified appetitive desire is related only to an unqualified object, which would be true if Plato holds the principle of relatives, the principle of opposites cannot be used to establish two subjects in the case of mental conflicts.

I have made an important assumption in this argument: I take for granted the identity of the object of an action and that in relation to which it is done. This, however, can be questioned.

IV. Object and Subjects

The strongest objection¹⁶ to my position, as I see it, will claim that the object of a willing action *completes* the action, while that in relation to which the action is done merely qualifies it. Thus, the object of an action is like the object of a transitive verb, which is required for the predicate phrase as a necessary component, while that in relation to which the action is done is like an adverb, which is an optional constituent. To speak of two actions as opposites is only possible when the action is viewed in tandem with its object, and it is a category mistake to then consider the object as that in relation to which the action is done. Sure, Plato then does not tell us whatever it is in relation to which an action is done; nevertheless, it cannot be the object of the action.

There in fact seems to be textual evidence for this: while in the principle of opposites relation is expressed in Greek by the preposition “πρός”, in the principle of relatives it is expressed through the objective genitive.¹⁷ Thus, it seems that when Socrates illustrates the principle of relatives by limiting drink itself to thirst itself, he is marking out the object to an action, and there is ground, based on the expressions Plato uses, to distinguish this from that in relation to which an action is done.

However, the plausibility of this objection is only apparent. At 437b, Socrates takes “assent and dissent, wanting to have something and rejecting it, taking

something and pushing it way” as pairs of mutual opposites.¹⁸ This identification is made, supposedly, so that when a pair on the list occurs, the principle of opposites can be used to make out two subjects. Applying the principle of opposites to distinguish subjects therefore involves a two-step procedural: 1) a pair of opposites is spotted, and 2) each of the three qualifying conditions is negated. Then, and only then, can two subjects be divided. Now, 1) can only be executed if the pair is really opposites, and not just possibly so, for in that case one cannot apply the principle of opposites in the first place.

I will grant that the objects which these actions are directed towards are not those in relation to which these actions are done. It is then clearly the case that the pairs listed at 437b are only *possibly* opposites if their objects remain indefinite: for example, assent to X can be opposite to dissent to Y, but this does not have to be the case. If assent is *always* opposite to dissent, X must be identical to Y. However, it has been shown in §3 that in the case of mental conflicts, the objects of the conflicting desires are *not* identical. Thus, in the case of mental conflicts, the pairs listed at 437b are only possibly opposites. In that case, one cannot validly apply the principle of opposites, for 1) has not been shown to be true. Thus, one still cannot validly infer, in the case of mental conflicts, that there are two subjects.

But couldn't it be the case that, *in fact*, when reason stops the thirsty person from drinking, it is reacting to the same object that thirst itself reacts to? This objection will now be considered.

V. Reason and the Good

In §4 a distinction was made between the object of an action and that in relation to which an action is done. I now make a further distinction: in contrast to 1) desiring X, there is, on one hand, 2) desiring $\neg X$, and there is, on the other hand, 3) rejecting X. One may then object that my argument rests on the identification of mental conflicts as 1) versus 2), when it should be 1) versus 3). If this is so, the pair of opposites, in the case of the thirsty person, for example, would have the same object, and thus really be opposites; and if the appropriate conditions are satisfied, Plato's claim that there are two subjects would be justified. However, can

¹⁶ For an objection which does not distinguish between “πρός” and “κατά”, and is thereby not as strong as the following, see Lorenz, *Brute* 24-25.

¹⁷ van Emde Boas et al., 371-372.

¹⁸ By “opposites”, Plato does not mean our notion of logical contradiction, but a more intuitive relation of one thing being the opposite to another. The intuitive sense of this relation probably has a spatial origin.

Plato be justified to suggest that in the case of mental conflicts what reason does fits the description of 3) instead of 2)?

Not if we accept his identification of reason as that which reasons about the better and the worse (441c). For, if this is so, reason's rejection is clearly not against the drink itself, for there is nothing bad about thirst *qua* thirst. Instead, it must be the element in addition to the drink itself which qualifies the *whole* drink as bad. 3) couldn't even be a consequence of 2), for in no way could reason go so far as to suggest that, just because drink combined with, say, poison is not good, *drink itself* is to be rejected.¹⁹ This is so with thirst and drink, and it is so too with the more general case of mental conflicts involving evaluative decisions.

Plato, therefore, could not have made a valid inference for the soul to contain multiple subjects, and this is so because the principle of opposites fundamentally conflicts with the principle of relatives, so that the first simply cannot carry out the conceptual duty which Plato needs it to. Nonetheless, the two external reasons which I presented at §2 for accepting this reading remain intact. If my account contradicts with some of Plato's own expressions and requirements, that is indeed unfortunate, but it only leaves my account less appealing. In my view, while external reasons could help us choose between two arguments which are equals otherwise, they should not have us accept an argument that does not make sense in the first place.²⁰

However, this is no reason to reject the tripartition entirely, and I now go on to discuss whether it could be taken as a theory which explains human actions, without making the theoretical commitment of dividing the soul into multiple subjects or agents.

VI. Normative and Motivating Reasons

A theory about the soul also makes human actions more intelligible by attributing them to certain sources.

A theory about human actions, on the other hand, engages with them directly by offering explanations. An action can be explained through its justification and/or motivation,²¹ both of which can be expressed as reasons for the action. Roughly speaking, a normative reason, which justifies an action, makes a case for why it is *good* for the agent to act as they do, while a motivating reason explains, from the agent's view, why they have acted as they do.²² Thus, if the tripartition is a theory that makes human actions more intelligible, one ought to ask, first, which type of reason Plato is thinking of. This question is not altogether straightforward to determine, however, since Plato does not distinguish these two types of reasons himself; still, it is an important distinction to make because his theory only works as one type of reason and not as the other.

First, the notions of a normative reason and a motivating reason require clarification; specifically, it is worth asking what their normative and motivating forces consist of. Now, if a reason is efficaciously motivating, it seems that it must thereby have an actively *determining* force. That is, its relation to the corresponding action is causal,²³ and if a motivating reason is posited, it should be followed by the action explained by it.²⁴ For I don't see how else to clarify the motivating nature of this type of reason. A normative reason, on the other hand, depends for its force on the conformity with or divergence from a norm that can be used for evaluation. Neither the reason itself nor the conformity or divergence is actively determining like a motivating reason; instead, what to do is up to the agent. Hence, the action doesn't have to follow when a normative reason is posited, and it makes no sense to speak of a normative reason as efficacious. If these are right, it follows that the same reason cannot be both normative and motivating at the same time.

What makes reading the tripartition as a classification of motivating reasons or forces so plausible is

¹⁹ This is where Reeve's defense fails: he thinks that thirst desires the drink itself and reason is against the "not-good drink", and it's impossible to divide in the object two such components (*Philosopher-Kings* 124-131). But manifestly reason only opposes the not-good-ness, not the not-good drink.

²⁰ Though in fact my position has two external arguments for it as well. One is that, right before he introduces the tripartition, Socrates warns about the inaccuracy of their method of investigation (435c-d). For the other reason, see note 24.

²¹ These do not exhaust all the possibilities of the ways an action can be made intelligible, but I think they are the ones most pertinent to my discussion.

²² I am aware that in the more recent debates in the philosophy of action, three types of reason have been distinguished, and the sense

I have given to motivating reasons is more properly the sense of explanatory reasons. I have chosen to go with the older locution because in the Plato scholarship, when the tripartition is treated as a theory of human action, it is usually described as a theory of motivation.

²³ This point is made by Donald Davidson in *Actions, Reasons, and Causes*.

²⁴ Note that nothing yet prevents the following action from being free or voluntary. Hence, John Cooper's seemingly paradoxical remark: "Plato's theory that there are three parts is, roughly, the theory that there are three psychological *determinants* of choice and *voluntary* action" (5).

probably in part due to the fact that Plato is certainly making a claim about mental states, and, under Davidson's influence, motivating reasons tend to be taken as mental states,²⁵ whereas normative reasons as facts. However, despite this agreement, Plato's tripartition simply doesn't work as a theory about motivating reasons. This point comes out most clearly in the case of mental conflicts, which I now go on to discuss.

VII. Mental Conflicts

Since Plato uses the phenomena of mental conflicts to first introduce the tripartition (439b-e, 439e-440b), it is appropriate that we focus on them for now. I will distinguish two models of mental conflicts. If a conception of mental conflict describes it as simply the interaction of determining forces, raging against each other inside a person, I will follow Christine Korsgaard and call it "the combative model". On the other hand, if a *person* deliberating and struggling to find the most appropriate course of action is being described, I call it a personal model. If Plato intends for his theory of action to be concerned with motivating reasons, his conception of mental conflicts would be combative, for each desire would then have a causal force on its own, acting on each other inside the person. If, however, Plato is thinking of normative reasons, the conflicts would be personal, for an agent is needed to deliberate what to do.

When reason and appetite are in conflict in the thirsty person, which model should be adopted? There are some textual evidence which show that Plato might have intended a combative model. At 439b, Socrates describes reason as "pulling" the thirsty soul and appetite as "driving" it; a bit further down, at 439d, appetitive desires are again described as "driving and dragging". Probably these have prompted Cooper to say that, "Socrates is conceiving reason as a *force* which works counter to appetite, pulling the agent back from what appetite pushes him towards" (7).²⁶ Reason and appetite would then be pulling the soul in different directions, as if in a tug of war, and whichever desire

wins out, its action then follows perforce. This counts in favor of the combative model, which means that reason and appetite are "motivational forces" in the mental conflict, as Cooper suggests, and it follows that reason and appetite characterize motivating reasons which are in a causal relation to their corresponding actions.

However, I think we should be wary of taking Plato too literally here. For, if reason is taken to be a force, alongside and against the appetitive forces, we would fall into some puzzles proposed by Korsgaard: "Are a person's actions merely the result of the play, or rather the combat, of these forces within her? How then would actions be different from blushes or twitches or even biological processes?" (2). That is, in the combative model of mental conflicts, the agent would resemble the trojan horse—a mere vessel—inside whom the forces decide the matter among themselves. Since the agent is not themselves a *force*, it would be difficult to see when and how the concept of a person acting could emerge at all. Now, if somehow a concept of the person is forced into this impersonal picture, Korsgaard continues, "what is the essence of this person, in whom reason and passion are both forces, neither of them identified with the person herself, and between which she is to choose? ... How—on what principle—can she possibly choose between them?" (2). That is, to the extent that the person is not identified with one of the forces, it is difficult to make sense of the nature of the person and their relation to the forces inside them.²⁷

Korsgaard's puzzles, I think, are not soluble unless reason is taken to be not just one among the many desires or forces. Indeed, when we experience mental conflicts ourselves, the very condition for the conflict seems to be our rational ability to *choose*;²⁸ there would be no conflict if it all comes down to forces acting against each other, for in that case the conflict resolves itself in an instant. The force of my desires is certainly a factor for my voluntary choice, but it is not a "determinant", as Cooper suggests.²⁹ According to

²⁵ Since the beginning of the 21st century, this thesis has been challenged, for example by Jonathan Dancy in *Practical Reality*.

²⁶ Cooper probably does not envision the parts as *merely* motivating reasons or motivating forces. He probably thinks that each is a real part of the soul, with a motivating force attributable to it, and in this respect his position is identical to, for example, Lorenz's.

²⁷ Couldn't a person simply *be* all or one of these forces? Why, it might be asked, is a separate conception of the person acting needed at all? It is needed because if a force is sufficient to count as a person acting, the soul would be divided into separate persons,

subjects, or agents, which has been shown, in §2-5, to be invalid. Neither can that point be now asserted on *this* ground, for there is no independent reason for accepting the claim that a force acting is to count as a person acting, except that it helps with dividing the soul into multiple subjects.

²⁸ The rational ability to choose clearly implicates a capacity to reason about the better and worse; Irwin's characterization of reason in §5 is thus confirmed.

²⁹ Cooper, 5.

Plato, it is by reason that we reflect on the desires and decide what to do with them (441e); even in cases where reason “succumbs”, the desires and their forces do not take over the choosing, but this ability of reason remains, to use a Hegelian expression, *overcome and preserved* (437c). Our own experiences confirm, then, that there is mental conflict to the extent that there is rational choice. It is natural to infer from this that, when we speak of a person as making a choice or acting voluntarily, we identify the person with their reason, since it is by reason that they do these things. It follows, then, as Reeve suggests, that, “reason...is the true human being rather than a mere (or proper) part of one.”³⁰

Thus, both from the problems that the combative model of mental conflicts leads to, and from our own experiences of mental conflicts, I think the personal model is to be preferred. If this is right, then Plato's theory works better if it is concerned with normative reasons, and not motivating ones. More specifically, reason, spirit, and appetite all represent desires, whose satisfaction counts as normative reasons, but reason, in particular, is also the capacity to deliberate about these normative reasons, and in respect to this it can be identified with the person.³¹ Thus, even though Plato does not himself make this distinction between motivating and normative reasons, it is more plausible to take his theory as about normative reasons.

In §6, I claimed that the same reason cannot be both normative and motivating at the same time. So far, I have tried to show only that Plato's theory does not work as motivating reasons *in the case of mental conflicts*. It is thus still an open question whether *in general* his theory could be taken as motivating reasons, or maybe sometimes as motivating reasons and sometimes as normative ones. I now turn to a discussion of the more general case.

VIII. The X-qua-X Language

To determine whether in general Plato's tripartition can be a classification of motivating reasons, I first assume the conclusion reached in §6, that motivating reasons stand in a causal relation to the corresponding

actions. According to Plato's principle of relatives, discussed in §3, unqualified things have relations with unqualified things, while qualified things have relations with qualified things (438a). This is to say, in the context of thirst, “thirst itself is, by nature, just for drink itself.” (439a). In general, then, if the three elements have motivating forces, each alone desires only the desideratum *simpliciter*.

Setting this aside, I will now try to make a hopefully uncontroversial remark about the *necessary* conditions for establishing a justifiable cause-and-effect relation. Now, a cause-and-effect relation is quite justifiably established when the cause is observable and prior, the effect is observable and contiguous with the cause, and there is constant conjunction of them.³² Whether these conditions are *sufficient* for establishing a cause-and-effect relation in a justifiable way, I do not wish to claim or investigate here; however, I think it will not be controversial to claim that, the skeleton of these conditions—namely, that 1) the cause is observable, 2) the effect is observable, and 3) there is constant conjunction of them—seems *necessary* to establish any cause-and-effect relation justifiably.

However, we do sometimes assert a cause-and-effect relation despite the fact that 1) is not satisfied. This is the case in the sciences when an unobservable theoretical term is posited to explain certain phenomena;³³ but here, their effects are always observable. In fact, the justification for positing unobservable causes precisely relies on the observable effects: we justify those theoretical terms because they can best make sense of the effects or go on to predict them. Thus, even if it is acceptable to sometimes violate 1), this violation is only justified instrumentally for the satisfaction of 2). It follows that, if both the cause *and* the effect are not observable, there is no adequate justification that can be given for establishing a cause-and-effect relation.

Going back to Plato, according to the theory of relatives, thirst itself, as a motivating force, causes wanting drink itself. Now, it is surely the case that thirst itself, and the appetitive part in general, is

³⁰ Reeve, *Blindness*, 83. It should be noted that Reeve reaches his conclusion primarily by reading 611b–612a and 518d–b, which seem to claim that the true nature of the soul is reason only, and because of this the soul is multiform. I count as in favor of my argument that the same conclusion can be reached by reading different parts of the text. See Reeve, *Blindness* 79–109.

³¹ Thus, while 1) a person may fail to be identified with reason's desire, 2) they cannot fail to be identified with reason's capacity of

deliberation. Cooper appreciates 1) without taking account of 2), and Korsgaard appreciates 2) but neglects 1). Cf. *Phaedrus* 246a–254e.

³² This is no more than Hume's observation of the cause as that which combines “three circumstances of contiguity, priority, and constant conjunction” (409).

³³ Physicists, for example, do this with concepts like “electron”, and biologists do this with “phenotypes”.

unobservable.³⁴ Therefore, if a cause-and-effect relation is to be justifiable, the effect, that is, the act, must be clearly observable. However, wherever is one to find this act, namely, wanting drink itself? For whoever has this desire perforce does not have the desire for a *good* drink, which violates the principle of relatives (438a); instead, such a person simply cannot and do not choose between a drink which is poisonous and a drink which is not, despite being aware of both as options. Keeping in mind that this is supposed to be a theory of human motivation, where can one ever observe from a human agent an act done regardless of these choices when they are aware of these choices and with the freedom to choose? Wouldn't even the most desperately thirsty person, to the extent that they can choose, always choose the non-poisonous drink?³⁵

But if in fact no one acts quite like this, then, as Robinson suggests, by employing "the sinister X-quax language",³⁶ Plato has shifted from the "historical, psychological, real problem" of human action to a semantic analysis.³⁷ This is why the act which is for the thing itself is never to be found anywhere in anyone. Thus, not only would what Plato posits as the motivating forces be unobservable, but so too would be the purported effects.

Now, it has been established that if neither the cause nor the effect is to be observed, the cause-and-effect relation cannot be justifiably established. Therefore, the cause-and-effect relation of the motivating reasons and the acts themselves cannot be established. It has been said in §6 that the motivating *forces* of these reasons just indicate that they are causes for actions that are explained by them. Therefore, without having established this cause-and-effect relation, the motivating forces cannot be established. Therefore, if the tripartition is concerned with motivating reasons in general, these motivating reasons would not actually

motivate, and a theory like this would surely not be very good. Therefore, also in the general case, Plato's division cannot be about motivating reasons.

IX. Conclusion

In this essay, I have discussed in what ways I find it implausible that Plato's division of the soul is either a theory of the constitution of the soul or a theory of motivation. I favor the reading that the three parts instead represent normative reasons, and reason, or the rational part, furthermore also represents the agent's capacity to deliberate. Thus, when Plato's thirsty man is going back and forth on whether he should drink something bad for him, it is not two parts of his soul or two motivating reasons that are in conflict; instead, his hesitation is due to the presence of two distinct normative reasons, which come into conflict³⁸ since what one of them justifies is the opposite of what is justified by the other.

In the end, this reading is not without its difficulties. For example, it would be difficult to make sense of Plato's conception of virtues in the *Republic* under this framework, and this only points to the further problem, emphasized by Moline, of how the city-soul analogy could work if the soul is not literally tripartite. Maybe this difficulty can be resolved, but I shall not work that out here. At any rate, I certainly do not pretend to have found even one final word on this topic; my task in this paper has consisted of rejecting two more or less orthodox views, which I believe to have been too little questioned for their own good, and in offering a new possibility, however far it is from being a coherent reading of the whole *Republic*.³⁹

³⁴ This is not to deny that we perceive our own thirst, but that we cannot perceive or observe "thirst itself", because while the former is an observable datum, the latter is a tool for explanation. As Robinson explains, "the psychologist is absolutely correct in saying 'You did that with your Id, or with your Reason, or whatever', because he means nothing by 'Id', 'Reason', etc., except the imaginary tool that does precisely that kind of thing" (47).

³⁵ If the case does not explicitly involve a choice between better and worse, then either it is a case of mental conflict discussed in §7, or, if there is no conflict, then certainly we also do not observe an act from thirst itself, for the act results from reason and appetite acting in concert.

³⁶ I take the expression "X-quax" as interchangeable with "X itself".

³⁷ Robinson, 41-42.

³⁸ Indeed, commentators have been remarkably comfortable to speak of two parts of the soul "in conflict" or of one part "overpowering" another, without clarifying exactly what this consists of. Should we imagine the soul engaging in pankration with itself? I think there is a clear and specific way to understand what is meant by Plato's expressions once we take the three elements to be normative reasons.

³⁹ The first draft of this paper was submitted for a tutorial with Prof. Roger Crisp, and I want to thank him for his comments and his various patient and always insightful discussions with me on Plato and ethics. I also want to thank Qiyuan Feng for reading a draft of this essay: his concerns were all worrying, but I could only respond to some of them. Lastly, I want to thank Cecilia Miller and Cara Cummings for their comments—it is their idea to bring up *Phaedrus*; I also want to thank the organizers at Prometheus for the wonderful conference.

Works Cited

- Adam, James. *The Republic of Plato: Edited with Critical Notes, Commentary and Appendices*. vol. 1, Cambridge University Press, 1905.
- Annas, Julia. *An Introduction to Plato's Republic*. Clarendon Press, 1981.
- Bobonich, Christopher. *Plato's Utopia Recast: His Later Ethics and Politics*. Oxford University Press, 2002.
- Burnyeat, M.F. "Plato on the Grammar of Perceiving." *Classical Quarterly*, vol. 26, no. 1, 1976, pp. 29–51.
- Cooper, John M. "Plato's Theory of Human Motivation." *History of Philosophy Quarterly*, vol. 1, no. 1, 1984, pp. 3–21.
- Crombie, I. M. *An Examination of Plato's Doctrines*. Vol. 1, Routledge & Kegan Paul, 1962.
- Dancy, Jonathan. *Practical Reality*. Oxford University Press, 2000.
- Davidson, Donald. "Actions, Reasons, and Causes." *Essays on Actions and Events*, Oxford University Press, Oxford, 2001, pp. 3–20.
- Hume, David. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J. Norton. Oxford University Press, 2006.
- Irwin, Terence. *Plato's Ethics*. Oxford University Press, 1995.
- Korsgaard, Christine M. "Self-Constitution in the Ethics of Plato and Kant." *The Journal of Ethics*, vol. 3, no. 1, 1999, pp. 1–29.
- Lorenz, Hendrik. "Plato on the Soul." *The Oxford Handbook of Plato*, 1st ed., edited by Gail Fine, Oxford University Press, 2008, pp. 243–266.
- . *The Brute within: Appetitive Desire in Plato and Aristotle*. Clarendon Press, 2006.
- Moline, Jon. "Plato on the Complexity of the Psyche." *Archiv Für Geschichte Der Philosophie*, vol. 60, no. 1, 1978, pp. 1–26.
- Moss, Jessica. "Appearances and Calculations: Plato's Division of the Soul." *Oxford Studies in Ancient Philosophy*, vol. 34, no. 1, 2008, pp. 35–67.
- Plato. *Complete Works*. Edited by John M. Cooper, Hackett, 1997.
- . *Republic*. Translated by C. D. C. Reeve, Hackett, 2004.
- Price, A. W. *Mental Conflict*. Routledge, 1995.
- Reeve, C. D. C. *Blindness and Reorientation: Problems in Plato's Republic*. Oxford University Press, 2013.
- . *Philosopher-Kings: the Argument of Plato's Republic*. Princeton University Press, 1988.
- Robinson, Richard. "Plato's Separation of Reason from Desire." *Phronesis*, vol. 16, no. 1–2, 1971, pp. 38–48.
- van Emde Boas, Evert, et al. *Cambridge Grammar of Classical Greek*. Cambridge University Press, 2019.
- Woods, M. "Plato's Division of the Soul." *Proceedings of the British Academy*, vol. 73, 1987, pp. 23–4.