

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXX · NUMBER 4

July 2025

ARTICLES

- 519 The Fair Share Theory of Conventional
Normativity
Adam Lovett
- 548 Introducing Discord
Mark Schroeder
- 572 Wronging Persons Through Photographs:
On Photographic Front Incursions
Macalester Bell
- 600 Desire Satisfaction and Temporal Well-Being:
Time for a New View
Frederick Choo
- 628 Delegitimizing Transphobic Views in Academia
Logan Mitchell

DISCUSSION

- 665 Gaslighting and Epistemic Competence
Cameron Domenico Kirk-Giannini

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXX · NUMBER 4

July 2025

ARTICLES

- 519 The Fair Share Theory of Conventional
Normativity
Adam Lovett
- 548 Introducing Discord
Mark Schroeder
- 572 Wronging Persons Through Photographs:
On Photographic Front Incursions
Macalester Bell
- 600 Desire Satisfaction and Temporal Well-Being:
Time for a New View
Frederick Choo
- 628 Delegitimizing Transphobic Views in Academia
Logan Mitchell

DISCUSSION

- 665 Gaslighting and Epistemic Competence
Cameron Domenico Kirk-Giannini

The *Journal of Ethics and Social Philosophy* (*JESP*) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge. Articles are typically published under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license, though authors can request a different Creative Commons license if one is required for funding purposes. Funding for the journal has been made possible through the generous commitment of the Division of Arts and Humanities at New York University Abu Dhabi.

JESP aspires to be the leading venue for the best new work in the fields that it covers, and it is governed by a correspondingly high editorial standard. The journal welcomes submissions of articles in any of these and related fields of research. The journal is interested in work in the history of ethics that bears directly on topics of contemporary interest, but does not consider articles of purely historical interest. It is the view of the editors that the journal's high standard does not preclude publishing work that is critical in nature, provided that it is constructive, well argued, current, and of sufficiently general interest. *JESP* also endorses and abides by the Barcelona Principles for a Globally Inclusive Philosophy, which seek to address the structural inequality between native and nonnative English speakers in academic philosophy.

JESP publishes articles, discussion notes, and occasional symposia. Articles normally do not exceed 12,000 words (including notes and references). *JESP* sometimes publishes longer articles, but submissions over 12,000 words are evaluated according to a proportionally higher standard. Discussion notes, which need not engage with work that was published in *JESP*, should not exceed 3,000 words (including notes and references). *JESP* does not publish book reviews.

Papers are published in PDF format at <https://www.jesp.org>. All published papers receive a permanent DOI and are archived both internally and externally.

Editors-in-Chief

Sarah Paul
Matthew Silverstein

Associate Editors

Rima Basu	Elinor Mason
Saba Bazargan-Forward	Simon Căbulea May
Brian Berkey	Tristram McPherson
Ben Bramble	Hille Paakkunainen
James Dreier	David Plunkett
Julia Driver	Sam Shpall
Alex Gregory	Kevin Toh
Sean Ingham	Mark van Roojen
Anthony Laden	Han van Wietmarschen
Coleen Macnamara	Jonathan Way
Vanessa Wills	

Symposia Editor

Errol Lord

Managing Editor

Chico Park

Copyeditor

Lisa Y. Gourd

Proofreader

Susan Wampler

Typesetter

Matthew Silverstein

Editorial Board

Elizabeth Anderson

David Brink

John Broome

Joshua Cohen

Jonathan Dancy

John Finnis

Leslie Green

Karen Jones

Frances Kamm

Will Kymlicka

Matthew Liao

Kasper Lippert-Rasmussen

Stephen Perry

Philip Pettit

Gerald Postema

Henry Richardson

Thomas M. Scanlon

Tamar Schapiro

David Schmidtz

Russ Shafer-Landau

Tommie Shelby

Sarah Stroud

Valerie Tiberius

Peter Vallentyne

Gary Watson

Kit Wellman

Susan Wolf

THE FAIR SHARE THEORY OF CONVENTIONAL NORMATIVITY

Adam Lovett

IN LONDON, people queue up at bus stops. Buses are frequent, so usually queuing is not necessary. But when everyone leaves work, sometimes there are too many people for all to get on the first bus. So Londoners form orderly queues, and the first to arrive at the bus stop is the first to board the bus. In this context, you act wrongly, morally speaking, if you jump the queue. In contrast, in Rome, people do not form orderly queues to get on public transport. When there are too many potential passengers for a bus in Rome, those who end up getting on are those quick enough to snag themselves a place. Do Romans act wrongly by not queuing? I doubt it. They are not obligated to queue because their social norms are different. There is no norm mandating queuing for buses in Rome.

Here is a different case. In New York City, diners generally tip 20 percent of the pretax bill when they eat out. They do this at cheap places and at fancy places, and they do it regardless of whether the server appears better-off than themselves. If you ever tip 10 or 15 percent on a meal, you have done something wrong, sometimes seriously so. In Tokyo, in contrast, diners do not tip. They may verbally thank the servers for a good meal, but tipping is neither required nor desired. You do nothing wrong when you do not tip in Tokyo restaurants. Again, the conventions in New York and Tokyo are simply different: no norm mandates tipping in Tokyo, and that is why one need not tip in Tokyo restaurants.

Consider one final case. Imagine you are an unmarried person in your thirties. If you are a Chinese person in China, your older relatives will likely give you a lot of advice about your situation. Your aunts and uncles will give you detailed input about your appearance, career, and comportment, all with the aim of getting you a spouse. Your parents will strongly encourage you, even pressure you, into settling down. In England, none of this usually happens. Your older relatives stay out of your business. Are the Chinese or English relatives acting wrongly in their respective situations? I am skeptical. Rather, familial norms in the two countries just differ. England has much stricter social rules protecting adult children from familial pressure on relationship issues. This is why the sort of advice elder relatives can give younger scions of their family is different.

These cases suggest that conventional, or social, norms have moral force. We are morally obligated to do some things that social norms dictate and would not be obligated to do those things in the absence of a social norm. Roughly, social norms are the standards that a group of people generally accepts for their behavior. Consider the rule forbidding you from eating with your elbows on the table or the rule mandating that you greet colleagues with a handshake: both are social norms in much of the Anglosphere. Or consider the norm that adult children can live with their parents until marriage. This is a norm in much of southern Europe. Such social norms sometimes correspond to legal norms. Traffic laws, for example, both impose legal obligations and (sometimes) match social norms. But social norms are often less formalized than legal norms. And such social norms sometimes directly correspond to independent moral norms. The norm forbidding murder is both a social norm and, independently, a moral norm. But social norms often do not correspond to independent moral norms. There is no general moral rule that you must stop your car at red lights, although in much of the world, there is a social norm demanding that you do. My aim in this paper is to explain why such social norms nonetheless often have moral force.

One might doubt the explanandum. One might doubt that social norms have any distinctive moral force. Perhaps we should comply with social norms only when we have independent moral reason to do as they direct. On this view, we have a duty to tip in New York city restaurants, for example, because we have a duty to prevent servers from falling into poverty. In this article, I largely just assume this skeptical view is wrong. For a start, that is because it is difficult to convincingly apply to specific cases. Consider tipping. At high-end restaurants in New York, the waiters earn much more than I do; still, it would be wrong for me not to tip. We cannot explain this via a duty to stop these particular servers from falling into poverty. More generally, the skeptical position does not explain the cultural variability of people's moral obligations. Londoners have a duty to queue at bus stops, but Romans do not; English parents have a duty to refrain from detailed commentary on the romantic status of their adult children, but Chinese parents do not. If these duties just flow from independent moral reasons, how do we explain why these reasons differ in different places? No explanation of this is obvious. Moral reality is not made of different stuff in Nanjing and Newcastle. These two points give us reason to look for a realist account of the moral significance of social norms.¹

The rest of the article embarks on this project. It is in part critical and in part constructive. Section 1 is critical. Laura Valentini, in a recent paper and book,

1 For a much more extended criticism of the skeptical view, see Valentini, *Morality and Socially Constructed Norms*, 53–60.

argues that we should obey social norms because doing so respects the agency of those who support the norms.² I raise two objections to this view that help establish some general features of social norms. In section 2, I outline an alternative view. On this alternative view, we should comply with social norms when and because we have certain collective obligations that those norms help us to discharge. Complying with norms is doing our fair share toward discharging these obligations. In section 3, I address the most concerning challenges to this theory. Overall, this provides good support for a fair share theory of the moral force of social, or conventional, normativity.

1. THE AGENCY RESPECT THEORY

The last twenty years have seen an efflorescence of work on the moral force of social norms. In this section, I focus critical attention on the view recently advanced by Laura Valentini.³ Partly, this is because her view is very interesting; partly, it is because evaluating it puts us in a good position to outline my positive view. Valentini thinks that following social norms is a way to respect people's agency. The first idea behind this view is that we have a "*pro tanto* moral obligation to respect other people's authentic and morally permissible commitments provided doing so is not too costly to us."⁴ Commitments are robust intentions. You might intend to be a good parent. This intention is robust if it is not merely transitory or fleeting: you retain the intention to be a good parent in a wide range of situations. These are the kinds of intention around which we orient our goals. Valentini thinks that to respect someone's agency, we have to respect these robust intentions.⁵ So we have a *pro tanto* moral obligation to help or at least not hinder the satisfaction of people's robust intentions. If you

2 See Valentini, "Respect for Persons and the Moral Force of Socially Constructed Norms" and *Morality and Socially Constructed Norms*. There is no substantial difference in position between the former paper and the latter book; the book just fleshes out the paper, so I will treat them as equally valid sources for Valentini's views.

3 For other views, see Gilbert, *A Theory of Political Obligation*; Marmor, *Social Conventions*; Miller, *The Moral Foundations of Social Institutions*; Melenovsky, "Promises, Practices, and Reciprocity" and "The Reasons to Follow Conventional Practices"; Owens, *Bound by Convention*; and Monti, "Are There Any Conventional Obligations?" Valentini addresses both Gilbert and Owens in *Morality and Socially Constructed Norms* (53–60). For criticism of Marmor, see Monti, "Are There Any Conventional Obligations?" 14–15. For criticism of Monti, see, in effect, the end of section 3.4 below. For discussion of both Miller and Melenovsky, see the end of section 2 below.

4 Valentini, *Morality and Socially Constructed Norms*, 10.

5 Valentini, "Respect for Persons and the Moral Force of Socially Constructed Norms," 390–92; and *Morality and Socially Constructed Norms*, 89–94.

are committed to being a good father, I have a *pro tanto* moral obligation to help you be a good father, provided certain conditions are met.

Valentini further adopts what she calls the *agential-investment* account of social norms.⁶ This account says that a social norm exists when a rule is widely and publicly accepted in a given context. To accept a rule, Valentini thinks, involves two things: first, “a cognitive state, namely a *belief*” that people ought to comply with the norm; and second, “an appropriate action-guiding attitude, namely a *commitment*”—the robust intention that people comply with it.⁷ The key part of this account is the second clause. This says that a social norm is in place only when people generally intend themselves and others to act a certain way. Consider, for example, the norm to queue or line up for buses. This is, according to Valentini, a norm only because many intend that all comply with the requirement to queue at the bus stop—they intend that this requirement functions as a “general standard of behavior.”⁸ Were people not to intend themselves and others to conform to this norm, then there simply would be no norm. The agential investment account links social norms closely to agency—it implies that social norms express the deep commitments of supporters of those norms.

These premises can straightforwardly explain the moral force of social norms. Norms express people’s commitments, and we have *pro tanto* obligations to respect people’s commitments. This just is to respect their agency. So we have *pro tanto* obligations to respect social norms. And that is to say we have a *pro tanto* obligation to follow social norms. The idea is, as Valentini puts it, that “the moral normativity of socially constructed norms stems from our duty to give people agency respect.”⁹ Now of course Valentini does not think we should always do as social norms dictate. For a start, our obligations here are merely *pro tanto*: they can be outweighed by other, weightier obligations. Additionally, it is only morally permissible, authentic commitments that we have *pro tanto* obligations to respect. I might have a commitment to upper-caste and lower-caste people not mixing, but this is a morally impermissible commitment, and so nobody has any *pro tanto* obligation to help me realize it.¹⁰ I might have a commitment only because I was manipulated into acquiring it. This is not, Valentini thinks, an authentic commitment, and so it does not

6 Valentini, “Respect for Persons and the Moral Force of Socially Constructed Norms,” 392–94; and *Morality and Socially Constructed Norms*, 22–30.

7 Valentini, *Morality and Socially Constructed Norms*, 23.

8 Valentini, *Morality and Socially Constructed Norms*, 26.

9 Valentini, *Morality and Socially Constructed Norms*, 82.

10 Valentini, *Morality and Socially Constructed Norms*, 91.

generate obligations.¹¹ Finally, if complying with someone else's commitment is too costly to our own agency—if, for example, it clashes too much with our existing commitments—Valentini thinks that we have no obligation to comply with it.¹² These conditions aside, though, the general idea is that we should obey social norms in order to respect the agency of supporters of those norms. She calls this the *agency-respect theory* of the moral force of social norms.

This is an attractive theory, but I do not think it is true. To begin with, I do not think we generally have a *pro tanto* obligation to comply with other people's commitments, even subject to Valentini's conditions. To see this, consider the following case. Suppose I am a Parisian. I form the authentic commitment that people in London stop queuing for the bus, that they decide who gets on the bus by playing rock-paper-scissors instead. This is a morally permissible commitment, or so it would seem. If it is permissible to commit to one way, queuing to allocate spots on public transport, it seems permissible to appeal to another—random chance—to allocate such spots. Yet my having this commitment gives Londoners no obligation, not even a *pro tanto* obligation, to play rock-paper-scissors to decide who gets on the bus. And nor do things change if I get a lot of my fellow Parisians to share my commitments. Even if the whole of France is committed to Londoners allocating seats on overburdened buses via a game of chance, Londoners have no *pro tanto* obligation to do so. And so we do not, generally speaking, have *pro tanto* obligations to comply with other people's morally permissible commitments.

Many cases show this. To take another example, imagine that I have strong commitments about what you wear when we meet. I intend you to wear red. It seems that this gives you no obligation whatsoever to wear red. My commitment in this case is morally irrelevant. Is that because my commitment is impermissible? That is doubtful. It would be impermissible to intend to force you to wear red, but I just intend you to wear red—I do not intend to force you to wear it. And it might be impermissible if my commitment to you wearing red was based on my belief that you had lower moral status than me or that I generally should be able to decide what you wear. But I might have no such belief. I might just think you look good in red. So I do not violate any rights against coercion or to equality. One might nonetheless think that we have a general right against others intending us to wear certain things. But this is false. We often permissibly intend others to dress in particular ways. If I invite you to my wedding, I might intend that you wear formal clothes. My having such an intention does not violate your rights. So the agency-respect view implies that

11 Valentini, *Morality and Socially Constructed Norms*, 90–91.

12 Valentini, *Morality and Socially Constructed Norms*, 91.

merely by intending that you wear red, I can give you a *pro tanto* obligation to wear red. This is false. So the agency-respect theory is false.

One might try to avoid this implication by claiming that complying with these commitments is too costly to generate a *pro tanto* obligation. But it is easy to formulate these counterexamples so that complying with the relevant commitments imposes very little burden on one's agency. Some Londoners do not care whether they queue for public transport. It would not be costly to these Londoners to change how they allocate seats on buses. Still, such Londoners have no obligation whatsoever to accept Parisians' commitments about such allocation. And you might not care how you dress. Indeed, you might be antecedently committed to wearing maroon. If so, it is not costly to you to wear red. Wearing red coheres rather than clashes with your existing commitments. Still, that I am committed to your wearing red gives you no obligation whatsoever to wear red. So such cases show that the agency-respect theory overstates the import of intentions. We simply do not generally have *pro tanto* obligations to comply with other people's commitments.

I want to consider another reply to this objection. One might suggest that in these cases, one does not entirely lack a *pro tanto* obligation to comply with other people's commitments. One merely has a very lightweight such obligation—an obligation that is very easily outweighed by other duties. On this view, our duty to comply with other people's commitments per se is lightweight. I do not think this accurately characterizes the import of other people's commitments about, for example, what we wear: these often generate no obligation, not just a lightweight one. But that aside, the problem with this reply is that, together with the agency-respect theory, it implies that we have only very lightweight obligations to comply with social norms. Yet often we have weighty obligations to comply with social norms. Think about the duty English older relatives have not to pressure their nephews to marry. This is no trifling thing. In England, if your aunt obsessively comments about your romantic life, she has done something seriously wrong. She owes you a heartfelt apology and should rectify her behavior. So this reply undercuts the capacity of the agency-respect theory to explain the phenomena.

What, then, is going on in the cases I have raised? I think the truth of the matter is that what commitments we have obligations to comply with is often determined by social norms themselves. You do not have any general obligation to comply with my commitments regarding your clothing choice because our norms determine that what you wear is part of your personal sphere. I do not get any say on it, except in special cases, as when I invite you to my wedding. And Parisians do not have any say on how Londoners allocate spots on buses because by the lights of our democratic norms, this decision is a matter for

Londoners alone. In contrast, I should not interfere with your commitment to your religion, for example, because our norms pick out that particular commitment as something to be respected. Which commitments matter and which do not is post- rather pre-institutional. It flows from the moral force of appropriate norms. If that is true, it augurs poorly for the agency-respect theory. More generally, it means that an appropriate theory of the moral force of social norms should help to explain how social norms shape the import of our commitments. In the next section, I advance such a theory.

Let us turn to a second objection to the agency-respect theory. This helps us get clear on what it is to accept a norm. The theory relies on the idea that when we accept a norm, we have a kind of agential investment—a commitment—to that norm. It says that to respect people's rational agency, we must respect these commitments. The thought, in Valentini's words, is that these commitments "are active: they are something that we author."¹³ To respect people's active authorship, one might naturally think, is at least a very good way of respecting their agency. The worry is that this is a deeply misleading picture of how most of us come to accept the norms we do. There is nothing very active about the way we usually come to accept norms mandating queuing or tipping or determining our familial rights and obligations. We do not reflect on these norms rationally and decide to endorse them. Rather, we are hardwired to seek out and internalize the norms that are extant in our social contexts. Norm following is an innate feature of our psychology, and our incorporation of norms is passive in a way akin to how we learn language or acquire aesthetic standards.¹⁴ When we see norm acquisition in this way, it is less clear that respecting them could be an especially important way to respect people's rational agency.

Let me say a little more about this picture of norm acquisition. The evidence for it starts from the fact that we begin to acquire norms when we are very young. Hannes Rakoczy, Felix Warneken, and Michael Tomasello report a nice experiment demonstrating this.¹⁵ They get a bunch of three-year-olds to come into their lab. In the lab is an adult playing with some objects: the adult puts some wooden pieces together to form a bat and uses it to push a block across a table into a gutter. The adult leaves without ever interacting with the children. The experimenter then comes into the lab with his hand puppet, Max. He has Max play with the objects differently to how the adult did. Specifically, Max lifts up the table so that the block slides into the gutter. The children respond by

13 Valentini, "Respect for Persons and the Moral Force of Socially Constructed Norms," 390.

14 For important sources of this picture, see Chudek and Henrich, "Culture-Gene Coevolution, Norm-Psychology and the Emergence of Human Prosociality"; and Henrich, *The Secret of Our Success*.

15 Rakoczy, Warneken, and Tomasello, "The Sources of Normativity."

telling Max off; he is doing it wrong. This suggests we evince norm-enforcing behavior as young children, long before we acquire the kind of practical agential faculties that warrant respect. So our norm-supporting behavior need not be explained by anything to do with these faculties. Even if norms do involve commitments, these are not the kind of reflective commitments that have deep moral significance.

We can underline this point by looking a bit more at the variation in norms across societies. People almost always endorse the norms they are raised with, at least in large part. British people think queuing is the appropriate way to allocate scarce goods. Americans think tipping 20 percent is obligatory. Chinese people see nothing wrong with interfering in the relationships of their younger relatives. This underlines the point that most people's absorption of norms is not a product of critical reflection. Critical reflection usually produces a lot of disagreement about normative matters. The history of ethics shows as much: when people think in a sustained and serious manner about what they ought to do, they come to very different conclusions. There is not much convergence in ethics; philosophers have very different views about moral issues. Within each society, though, there is a lot of convergence on which norms to endorse. People simply adopt the norms extant in their society. All that suggests that a mechanism besides critical reflection is leading to their adoption of norms.

What is this mechanism? A view recently championed by Joseph Henrich is that we are simply innately disposed to imitate other people, especially prestigious people.¹⁶ We do this primarily by working out what norms exist in our environment and internalizing them. Here, to internalize a norm just is to be committed to both complying with it and sanctioning others for noncompliance. There is an evolutionary explanation for why we do this. Knowledge is extremely important in human societies: it is vitally important to know how to make fire, cook casava, or find edible plants. But the relevant kind of knowledge is hard to verbally communicate. It is complicated and often tacit. So the best way to utilize such knowledge is simply by imitating other people, especially successful (and thereby prestigious) people. And we do this by internalizing norms. The key point is that the acquisition of norms here is arational rather than rational. There is a sensible causal explanation for why we are disposed to internalize norms, but we do not think through these issues when we internalize norms. (Children surely do not.) We are simply strongly disposed to internalize norms.

16 Henrich, *The Secret of Our Success*.

Why does all this matter? Because commitments should be taken most seriously, I think, when they are reflective. They should be taken most seriously when those who have them have thought carefully about them and, upon detailed consideration, decided to endorse them. Perhaps unreflective, unconsidered commitments also have some moral force. But they reflect people's actual exercise of rational agency to a far lesser degree than do other commitments and so have far less force. Yet our commitment to norms, I believe, is usually unreflective and unconsidered. British people do not reflect on whether queuing is the best way to allocate seats on a bus; they are arationally committed to queuing. If that is so, the commitments identified by the agency-respect theory are ones we should not take too seriously. They are not ones that reflect the actual exercise of rational agency, and so respecting agency does not require much of us with regard to such commitments. I doubt such morally weak commitments are sufficient to explain the force of social norms. Social norms are often morally weighty, but any commitments underlying them are morally lightweight.

Let us look at just one worry about this position. Does it make too many of our commitments morally lightweight? Take, for example, religious commitments. One might doubt that all Catholics, for instance, have carefully reflected on their commitments. But religious commitments are nonetheless morally heavyweight. The reply to this point is that the moral weight of a commitment need not hinge on the nature of that particular commitment. Rather, on the theory to be outlined in the next section, it hinges on the moral weight of commitments of that kind. Thus, some people's religious commitments may be unconsidered and unreflective—in themselves, these are not worthy of much respect. But many people have deeply reflective, considered religious commitments. So we should institute and maintain social norms that mandate we respect all religious commitments. The unreflective commitments get protected as a side effect of such norms. I spell out this position in more detail in the next section. The upshot of it, though, is that the position that I have just outlined does not make too many of our commitments into morally lightweight ones.

That concludes my discussion of Valentini's agency-respect theory. I think this theory is initially attractive but ultimately untenable. In showing that, I have made several crucial suggestions about social norms. Most importantly, I have claimed that norm following is a deep-seated, unreflective part of human psychology. And I have suggested the norms alone often generate weighty moral reasons, that they shape what commitments we have a duty to respect, and that we have a duty to respect some—but only some—unreflective commitments. In the next section, I advance a positive proposal that does justice to these claims.

2. THE FAIR SHARE THEORY

My positive proposal descends from the Rawlsian idea that we can explain many obligations in terms of a duty to do our fair share.¹⁷ Rawls understands certain obligations such as promissory obligations in terms of a duty to contribute our fair share to the maintenance of cooperative schemes from which we benefit. This is a kind of duty of reciprocity. The idea in the promissory case is that we benefit from the practice of promising, and keeping our promises is making a fair return for that benefit. These specific ideas do not easily generalize to social norms. The chief problem is that often we do not benefit from social norms that bind us.¹⁸ Consider the tipping norm. It is primarily American wait-staff who benefit from the norm that 20 percent be tipped on every meal. I have never been a waiter. I do not benefit from this norm, or at least do not benefit much.¹⁹ Still, I am bound by the norm. I do something seriously wrong if I leave a New York restaurant without tipping. So my positive proposal distantly descends from Rawls's idea, but the details are very different.

I propose that we start with the claim that we have many collective obligations.²⁰ Collective obligations are duties that we have not simply as individuals but as a plurality. Imagine, for example, that a stranger is stuck below a log that needs five people to lift it. You do not have an individual obligation to lift it (you cannot), but if there are exactly five people in the area, you together have a collective obligation to lift the log. Your contribution makes a difference to freeing the stranger, but that is not so in all cases of collective obligations. Imagine, for instance, that there are one thousand wounded men lying in the desert suffering from intense thirst. You are at a nearby banquet with an equal number of people. Each banqueter has a pint of water. You could pour your pint into a water cart, which will be driven to the desert where the water will be equally distributed among the wounded. But you pouring your pint will have no perceptible effect on slaking anyone's thirst: one-thousandth of a pint

17 For this view, see Rawls, *A Theory of Justice*, sec. 52. Rawls himself was influenced by Hart, "Are There Any Natural Rights?" 185.

18 There are also other problems brought out by discussion of this theory applied to political obligations. For these problems, see Simmons, *Moral Principles and Political Obligations*, 101–42; and Huemer, *The Problem of Political Authority*, 86–93. For a more extended critique, see Monti, "Are There Any Conventional Obligations?" 12–14.

19 Do I benefit because the norm yields good service? Well, we do not need tipping to get good service—service in Japan is excellent.

20 The positive proposal I advance here has some similarity to that Wellman advances in "Samaritanism and the Duty to Obey the Law." The differences are that Wellman is interested in duties to obey the law rather than social norms more generally, and relatedly, he does not mention the empirical work on which I rely.

of water is imperceptible. Still, you together with the other banqueters are *collectively* obligated to fill the cart with water. If you do not fill up the cart, you collectively have done wrong.²¹

These are stylized cases. But there are many familiar examples of collective obligations. Consider, for example, our duty to avert very serious global warming. This is not a duty any individual has. None of us can individually avert serious global warming. Rather, it is a duty we have collectively as a society, or as the whole of humanity. Equally, we have collective obligations to ensure goods are fairly distributed in our society, that people are paid sufficiently for their work, that everyone has a decent chance at authoring their own life, that children are well cared for, and much else besides. We are not just individually obliged to ensure these things; we are collectively obliged to do them. Such collective obligations are not simply aggregations of the individual duties of the members of this plurality. It is not that we are each individually obligated to avert global warming. Instead, collective obligations are primarily duties that apply to collectives. We should think of them as obligations that fundamentally attach to pluralities of people.

When exactly do we have collective obligations? Generally, whenever there is a very morally serious goal that we must coordinate to achieve, we seem to have a collective obligation to achieve the goal. I have two remarks about this. First, coordination matters. If each of us could easily lift the log on our own, freeing the stranger trapped under it, then I doubt we would have a collective obligation to lift the log. Each of us would simply be individually obligated to lift it. For a goal to generate collective obligations, coordination must be necessary for achieving it. It must be necessary that people respond to and rely on the actions of others. Second, the goals that generate collective obligations are varied. These are not just the goals of ensuring that certain goods or burdens are distributed fairly. They include the goal of ensuring certain goods (familial relationships, say) exist at all, the goal of saving people's lives, the goal of protecting rights. I do not have a general account of what makes a goal morally serious enough to generate a collective obligations. This is a hard question that I will leave open. But for now, the important point is simply that collective obligations are common and can be in the service of multifarious goals.

Let us move on to a key further observation. It is very hard to discharge our collective obligations. Consider the duty to ensure that goods are fairly distributed in our society. There are two things that makes this very difficult to discharge. First, there is a motivational problem. We cannot reach into everyone's heads and ensure they give appropriate weight to fair distribution on each

21 This case is discussed in more detail by Parfit, *Reasons and Persons*, 75–86.

occasion they deliberate. We are not capable of influencing people's practical deliberations in this fine-grained way. The problem is exacerbated by the fact that people have strong countervailing motivations. We are selfish: we are disposed to give ourselves more than our fair share. And we are partial: we give our friends and family members more than their fair share too. This means merely trying to persuade people to give appropriate weight to fair distribution is of limited efficacy. Second, there is an epistemic problem. To satisfy our collective obligations, people have to know what others are doing. Whether giving you some good on some occasion is fair depends on what you get on other occasions. But it is very difficult for me to know what you get on other occasions, and even if I did know, it would be very difficult for me to work out whether it was fair relative to other people's allocation. It is difficult to predict the behavior of other people.

These problems present enormous barriers to discharging our collective obligations. Fortunately, there is a solution to both problems. In our discussion of the agency-respect theory, I staked out a particular account of the psychology of social norms. On this account, we human beings are norm-following creatures. We are innately and strongly disposed to evaluate what norms hold in our social group and then to adopt those norms. We are disposed to take those norms as a standard for our own and others' behavior. This, as discussed above, is an arational process, and our disposition to follow norms often overwhelms our other motivations; it often overwhelms our selfishness, our partiality for friends and family members, our apathy. British people, for example, feel terrible about jumping a queue, even when they are about to miss their flight; people refrain from lying or stealing or marrying distant cousins, because their social norms forbid it. Norms have a pervasive influence on our behaviors.

That means, I suggest, that often the only way to discharge our collective obligations is for us to collectively support—to set up and maintain—norms that, if complied with, would discharge the obligations. This solves the motivational problem because norm following is such a powerful human drive. Setting up norms is an indirect way of reaching inside people's heads in order to ensure they deliberate appropriately. And as a result, it solves the epistemic problem. We can predict others will comply with the norm because human beings are norm-following animals. To see this solution at work, reconsider the duty to ensure goods are fairly distributed in our society. One way to ensure certain goods such as seats on public transport are fairly distributed is by queuing. It is fair to equalize the time everyone spends waiting for their bus, and a queuing system tends to do that. Those at the back of the queue will on average end up waiting about as long for their bus as those at the front have already waited. So the most reliable way for us, *qua* collective, to discharge our collective obligations is to support social norms that, when followed, will discharge them.

Thus, we have collective obligations to support social norms that, if complied with, help us to discharge our collective obligations. This inference relies on a means-end principle: if we are collectively obligated to bring about *X*, and bringing about *Y* is the most reliably means of bringing about *X*, we are collectively obligated to bring about *Y*. The kind of principle is plausible for individual obligations: if *Y* is the most reliable way of discharging your individual obligations, you are usually obligated to *Y*.²² The principle is equally plausible for collective obligations. Consider the stranger stuck under a log. The most reliable way to lift the log off the stranger might be for the five of you to make a plan about how to lift it. So you have a collective obligation to make such a plan. The collective obligation derives from the fact that plan making helps to discharge your collective obligations. Setting up or maintaining certain social norms is the most reliable and perhaps the necessary means to discharging certain of our collective obligations. So we have a collective obligation to set up such norms or to maintain them if they are already in place.

Why does that matter to our individual obligations? Here, we need a further moral claim—and this connects the present view with the Rawlsian idea. The claim is that when we have a collective obligation to do something, we are individually obliged to contribute our fair share toward that thing.²³ Think again about global warming: we are collectively obliged to avert very serious global warming, so we are individually obliged to contribute our fair share toward averting such warming. Concretely, we are obliged to reduce our consumption, to offset our emissions, or to install solar panels. More generally, I think we are obliged to make the kind of efficacious sacrifices such that, if everyone made these sacrifices, we would discharge the collective obligation.²⁴ It is a sacrifice to not fly transatlantically or to pay money to offset your household emissions. But these sacrifices contribute to averting serious global warming, and if everyone made such sacrifices, then we would succeed in averting serious global warming. And so making such sacrifices is a way of doing your fair share toward contributing to this goal.

22 In some cases, the most reliable means to fulfilling some obligation might be much more personally costly than some slightly less reliable means. These cases may prove exceptions to the principle. But such cases generally do not matter to our discussion of social norms.

23 One recent source for this claim is Dietz, "What We Together Ought to Do." Wellman ("Samaritanism and the Duty to Obey the Law") and Maskivker (*The Duty to Vote*) also endorse a claim of this sort.

24 For this kind of view, see Murphy, *Moral Demands in Nonideal Theory*. For a different approach, see Tosi, "Rethinking the Principle of Fair Play." Tosi's approach provides distinctive resources for addressing the problems of discretion discussed in section 3.4.

That completes the fair share theory of conventional obligations. The theory is rooted in the causal centrality of social norms to human behavior: we are norm-following creatures. Human beings are the animal that adopts and follows social norms. This centrality means that setting up and maintaining good social norms is the most reliable means for us to discharge our collective obligations. Thus, when we have such norms, we are collectively obligated to maintain them. So we have an individual obligation to do our fair share toward maintaining good social norms. We each should make the sacrifices such that if everyone made similar sacrifices, the norm would be maintained. In most cases, what those sacrifices are is fairly clear: we should internalize the norm. We take it as a standard for our own and other people's behavior. That means we will comply with it and sanction others for noncompliance. If everyone did this, the norm would endure. Norms, then, have moral force when and because internalizing them is doing our fair share toward discharging our collective obligations. We have many collective obligations that norms help us to fulfill, and internalizing those norms is our way of contributing to such obligations.

Let us see how this theory applies to our earlier cases. Start with queuing. Here, I think the relevant collective obligation is to fairly distribute public goods. In London, maintaining the queuing norm is the most effective means of discharging this obligation. So we have a collective obligation to maintain the queuing norm. We should each contribute our fair share to maintaining this norm. So we should each internalize and thereby comply with it. So Londoners have an individual obligation to queue up at buses in order to contribute to discharging their collective obligations. In Rome, in contrast, there is no such norm and so no such individual obligation. Perhaps Romans should set up a queuing norm, but in its absence, they have no duty to queue. Now turn to tipping. Here our collective obligation is to ensure that workers are paid sufficiently for their work. In the United States, the tipping norm ensures that waitstaff are paid sufficiently for their work. New Yorkers have a duty to do their fair share in maintaining this norm, and thus they should tip. In Japan there is no such norm, and so people are not obligated to tip.

Familial norms are a little more complicated. These implicate, I think, two of our collective obligations. On the one hand, we have a duty to help people author their own lives, to help them ensure that their important choices manifest their own values, intentions, and commitments. On the other hand, we have a duty to ensure people can have close and rewarding family relationships. This includes relationships in which they can give advice to family members freely. There are plausibly many equally good ways of balancing these goals. The English norms that constrain older relatives from giving certain kinds of advice to younger scions tends to more promote the first goal. Chinese norms

that permit such advice tend to more promote the second. Both are, I think, equally fitting ways to discharge our multiple, conflicting, collective obligations. If so, then one has a fair-share duty to comply with the norms in the society one finds oneself in. So neither Chinese nor English people do anything wrong in complying with their respective norms. Generally, the moral force of social norms comes from a duty to contribute our fair share to discharging a collective obligation.

At this juncture, the fair share theory should be clear. So let us tie up some loose ends. First, in the previous section, I promised that the fair share theory would explain why, although unreflective and unconsidered commitments lack intrinsic import, they should often still be respected. Let us fulfill the promise. We start with the idea that we have a collective obligation to respect people's authentic and reflective commitments. This is a corollary of our obligation to help people author their own lives. The key point is then that for many kinds of commitments, we do not effectively discharge this obligation by setting up a norm that lets individuals judge which commitments of this kind are authentic and reflective and which are not. Third parties do this badly; they do not have the information to tell whether a commitment is authentic and reflective, and their evaluation is predictably biased. We are much more likely to think a commitment is reflective if we share it. So when certain kinds of commitment are often authentic and reflective, such as religious commitments, we should support norms that mandate respect of all such commitments. So we should respect even unreflective and unconsidered such commitments. Doing so contributes to norms that effectively discharge our collective obligation. More specifically, we should respect religious commitments in general, even the unreflective ones, because many religious commitments are reflective.

This connects to a second point. Previously, I suggested that what commitments we have a duty to respect often depend on the structure of our social norms. I have a duty to respect Londoners' commitments about how to allocate public transport in London but not those of Parisians. You have a duty to respect my commitment that you wear a suit at my wedding but not my commitment that you wear red all the time. Generally, this is because our duties to respect commitments are filtered through social norms. We should respect those commitments that good norms—norms that help us to discharge our collective obligations—tell us to respect. The norm that gives you control over what people wear at your wedding (within reason) is such a norm. There is no norm that gives you control over what people wear all the time, and if there was, it would not be a good norm. It would not help us generally discharge our collective obligation to ensure others author their own lives. And so the fair share theory explains why our autonomy rights so often seem post- rather than

pre-institutional. I believe (controversially) that much of the structure of our autonomy rights is determined by our social norms.

Besides these loose ends, I have so far left some other questions about the fair share theory open. Let us explore some of these issues. To begin with, there is an important question of when a social norm is sufficiently good at helping us to discharge our collective obligations as to generate obligations. One could, in answering this, say that each collective obligation is defined by a goal: the goal of fair distribution, reciprocity, self-authorship, etc. Then one could say that a norm generates individual duties as long as it brings us above some threshold of achieving that goal. This latter idea, however, seems implausible. The problem lies in defining a threshold; any threshold here seems arbitrary. A better view then seems to be a simple counterfactual view: when evaluating a norm, we see whether the relevant goal would be better achieved were the norm not to exist. Here we evaluate what other norms would take its place were it not to exist (and the answer might be none). What matters here is the likely replacement for the norm were it to disappear. If, in this situation, the relevant goal would be less well achieved than it actually is, the norm's existence brings us closer to achieving the goal. And so we should do our fair share to uphold the norm.

A second important question concerns what the ideas behind the fair share theory imply when we do not have relevant norms in place. To take an extreme case, imagine we live in Hobbes's state of nature and have no shared norms at all. In this case, we have a collective obligation to set up shared norms. So we have an individual duty to do our fair share toward setting up norms. But what that requires is much more varied than when we already have norms in place. For a start, we need to evaluate what norms to try to set up. That involves evaluating both how good different norms are and how feasible they are to establish. We then need to work out the best way for us to contribute toward setting them up. This might be some mixture of activism and conformity to the norm before it is prevalent. Evidently, what to do in such cases is much more multifarious and context dependent than what to do when we already have a good norm in place. Fortunately, to explain the moral force of actually existing norms we need only to consider the simpler case.

Finally, it is worth comparing the fair share theory of conventional obligations to some related theories in the literature. I have already compared it to a Rawlsian theory—the crucial difference is that on the fair share theory, one need not benefit from a convention in order to have a duty to uphold with it. C. M. Melenovsky has advanced a view with some similarity to the Rawlsian picture.²⁵ Melenovsky attempts to ground the moral force of social norms in a

25 Melenovsky, "Promises, Practices, and Reciprocity."

certain kind of reciprocity. His view is that as a *sui generis* principle, if one follows the rules of a sufficiently just social practice, then one has a claim on others to follow those rules.²⁶ The practice has to be “sufficiently just” in order to generate such claim—this is why those who participate in, e.g., chattel slavery do not have a claim on others to participate. But any practice that exceeds this bar generates the claims. So, for example, those who follow London’s queuing norms have a claim on other people to follow those norms. Both the Rawlsian theory and Melenovsky’s are based on some sort of reciprocity. Melenovsky, though, focuses on a claim that contributors to a norm allegedly enjoy rather than on a duty that beneficiaries of the norm are constrained by.

How we understand Melenovsky’s view turns on how we understand what it means to be sufficiently just. One might think that a practice is sufficiently just if and only if—and because—it helps us to discharge some collective obligation. On this understanding, Melenovsky’s *sui generis* principle can be straightforwardly thought of as identifying a claim on others to do their fair share in discharging collective obligations. So this view collapses into the fair share theory I have advanced. But one might more distinctively think a practice is sufficiently just when it does not violate people’s rights—such as in the example of chattel slavery—or otherwise impose extreme demands on people. Unfortunately, this view has difficulty dealing with relatively pointless social practices. Think, for example, of the rule that one cannot take liquids onto flights in containers of more than one hundred milliliters. I doubt this does much to improve the safety of airline passengers. Yet this rule does not violate anyone’s rights or impose onerous demands on anyone. So the view under discussion implies that those who follow this rule have a claim on others to also follow the rule. But this seems to me implausible—completely pointless social rules lack moral force. So I doubt that either of these versions of Melenovsky’s view provides a plausible alternative to the fair share theory.²⁷

One feature of both the Rawlsian view and Melenovsky’s view is that collective obligations play no official role. So let us compare the fair share theory of conventional obligations to a view in which collective obligations play a prominent role. Seamus Miller advances a theory to explain, *inter alia*, why

26 Melenovsky, “Promises, Practices, and Reciprocity,” 112.

27 For a later view, see Melenovsky, “The Reasons to Follow Conventional Practices.” Here, Melenovsky adopts a pluralistic account of our reasons to follow social norms: he suggests there are many different kinds of reasons to follow social norms, of which the consideration in the text is a part. I do not disagree with this overall pluralistic approach. Still, I think it is useful to look for the reason that applies in the widest array of cases. This is in some sense the most fundamental reason to follow social norms, and this is the target of the fair share theory of conventional normativity.

the members of organizations have distinctive duties.²⁸ He thinks that organizations help us to produce goods that we are collectively obligated to produce. Organizations that discharge these collective obligations are made up of different roles (think of teacher in a school or police officer in a police force). And, Miller thinks, people who occupy these roles are required to comply with the norms constituting their roles.²⁹ Both Miller's view and the fair share theory invoke collective obligations, but their target explanandum is different. Miller's target is why organizational norms have moral force for their members. The target of the fair share theory is why social norms more generally have moral force. So the target of the latter is much broader than the former.

Still, perhaps we can modify Miller's theory to explain the broader explanandum. We could think of society itself as an organization that discharges certain collective obligations, of a "member of society" as a kind of role, and of norms like the queuing norm as constitutive of this role. This brings us to the more important point of contrast between the two views. Miller does not really say anything about *why* the individuals who occupy a role are obligated to obey its constitutive norms. He says it is not because they promise to comply with those duties or because their embodying the role creates reasonable expectations for others.³⁰ But he does not give a positive explanation for it: role responsibilities are an unexplained explanans.³¹ The fair share theory explains this explanans. Those who occupy the roles in good organizations have a duty to do their fair share to contribute to the organization's proper functioning. And usually the most effective way for them to do this is by fulfilling the demands of their roles. So the fair share theory provides a deeper explanation of the duties of individual people than does a role responsibilities view.

That is the fair share theory of the moral force of social norms. We have now seen the ideas that motivate it and how it is situated in contrast to other views. Let us thus turn to addressing the most serious objections to it.

28 Miller, *The Moral Foundations of Social Institutions*.

29 Miller, *The Moral Foundations of Social Institutions*, 80.

30 Miller, *The Moral Foundations of Social Institutions*, 78–79.

31 A reviewer suggests that this is addressed in Miller, *The Moral Foundations of Social Institutions*, ch. 4. But this chapter, as I read it, mainly seeks to explain when collectives are morally responsible. Miller is not trying to explain why the moral obligations of individuals can derive from roles and so does not give a deep explanation of individuals' obligations.

3. CHALLENGES

3.1. *Directed Obligations*

The first serious challenge to the fair share theory is how it captures the directness of our duty to obey social norms. When you jump a queue, you do not equally wrong everyone who benefits from the queuing norm: you wrong the people in the queue especially. It is to them you owe an apology; it is them you must compensate for your wrongdoing. And they have special standing to resent and criticize you. Yet on the fair share theory, it is not clear why that would be. For on the theory, jumping a queue is wrong because it is not doing your fair share to uphold norms that help to discharge our collective obligations. But your duty to do your fair share to uphold such norms is owed to all others with the collective obligation, not to those in the queue especially. So it is not clear why those in the queue are wronged especially by your queue jumping.³²

This is also a problem for the agency respect theory. On this theory, queue jumping wrongs those who are committed to the queuing norm. But those outside the queue might be just as committed to the norm as those in it, so it is not clear why you wrong those in the queue especially. In response to this problem, Valentini suggests that “some social norms confer rights as control on individuals.”³³ If you are in a queue, the queuing norm grants you the ability to control whether I may go in front of you or not. The norm does not simply forbid me from going in front of you. Rather, it says that you must not jump the queue unless everyone in it says you can. Thus, people in the queue can impose obligations on you. Moreover, Valentini believes that all it is to be the recipient of a directed obligation, in an important sense, is to have the kinds of control rights that social norms sometimes confer on individuals.³⁴ The idea is that I wrong you when I do something impermissible that you had the power to forbid. So we wrong people by jumping a queue.

This strategy seems unsatisfactory. I doubt that violating such control rights is sufficient for wronging someone in the relevant sense. To see why, notice that we can give people control rights artificially. Suppose you live in a brutal autocracy, and the dictator says he will start killing innocent people unless you obey his son. You should do what the son says. The son has the power to impose obligations on you: when the son tells you to do something, you should do it. Yet you do not wrong the son by spurning his commands. You do not owe

32 This is David Owens's main objection to the kind of view I endorse. See Owens, *Bound by Convention*, 29–30.

33 Valentini, *Morality and Socially Constructed Norms*, 215.

34 Valentini, *Morality and Socially Constructed Norms*, 123–26.

the son compensation or an apology for disobeying him, nor can he fittingly resent you for such disobedience. So merely violating someone's "rights as control" does not wrong them in the sense that we wrong the victims of our norm violations. I think this worry generalizes to any version of Valentini's strategy. At root, Valentini aims to identify a particular normative position and say that whether an impermissible act wrongs you depends on whether you are in this position. But when we create normative positions artificially, those in such positions (e.g., the dictator's son) are not wronged by relevant impermissible actions. So I doubt any story of this kind captures the kind of directed obligations operative in norm violations.

Fortunately, there is another, better story available (to both the fair share theory and the agency respect theory). This other story hinges on the observation that when you violate a norm, the *harm* of your norm violation does not fall on everyone equally. It falls especially on the victims of your violation. When you violate the queuing norm, the harm of your violation falls on the people already in line. It is they who have to wait longer for a seat on the bus. The people most harmed by your wrongdoing are those waiting in line, not all those who benefit from the queuing norm. I propose to use this fact to explain the directedness of conventional obligations.

To do this, we need a substantive claim about directed duties. I believe we have weighty directed duties not to foreseeably harm people by our wrongdoing. We owe it to people not to foreseeably set back their interests as a result of our wrongful actions.³⁵ To provide some evidence for this claim, consider recklessness. Suppose you drive home drunk. You wrong everyone on your route home by imposing a risk on them. Now suppose additionally that because you are drunk, you lose control of your car and crash into someone, injuring them. You harm the person you injure. By harming them, you wrong them more seriously than those you merely imposed unrealized risks on. Yet it is not wrong in itself to lose control of your car and crash; if you had crashed because of ice or fog or bad luck, you would not have wronged the person you injured. You wrong this person because your crash stemmed foreseeably from your own prior wrongdoing—your decision to impose risks on people. And you owe it to them not to foreseeably harm them by your wrongdoing.

This allows us to explain why the perpetrator of a norm violation wrongs the victim especially. The harms of their wrongdoing foreseeably befall the victim, and they have a weighty moral obligation not to foreseeably harm people by their wrongdoing. This is a directed obligation, so the perpetrator has violated

35 For a more extensive defense of this kind of view, see Cornell, "Wrongs, Rights, and Third Parties."

a weighty directed duty to the victim. This is why they owe it to the victim to apologize to them. They must apologize for harming them via their wrongdoing. And it is why the victim can resent them; they can resent that the perpetrator's wrongdoing has led to a harm befalling them and so wronged them. As I have said, this view is open to both the agency respect and the fair share theories of conventional normativity. I believe it captures why social norms generate especially directed obligations.

Let me make two remarks about this view. First, I do not endorse the claim that whenever your actions cause some harm to befall someone, then you wrong that person. Causation is too promiscuous for this to be plausible. Suppose my wrongfully lying to someone causes them to release some butterflies in Brazil. One of these butterflies flaps its wings, and that causes a tornado in Texas. I harm some Texans, but I do not wrong them. Instead, the principle I endorse says that it is only when a harm is the *foreseeable* consequence of your wrongful actions that you wrong those who suffer the harm. Something is a foreseeable consequence of my actions, I think, when I have sufficiently good evidence that it will result from my actions. I could not, by this standard, foresee that my lie would cause a tornado in Texas. But when I jump in front of someone in a bus queue, it is very easy to foresee that it will set back one of their interests—specifically, their interest in getting on the bus as soon as possible. So it is specifically foreseeable harms that generate directed obligations out of undirected wrongs.

Second, I want to address a worry about the view. One might think that some violations of social norms are harmless but still wrong the victim especially. Suppose we are at a campsite, and I take your barbeque gear without your permission when you are away for the day.³⁶ I put it back before you return, and you never know about the theft. I have wronged you especially, but where is the harm? You might think my theft is a harmless wrongdoing, that it leaves you no worse off. But that thought rests on too narrow a conception of harm. Harms are not only what give you bad feeling or waste your time; harms include things that set back your autonomy interests. These are interests in controlling parts of the world, deciding what happens in those parts of the world.³⁷ My theft impairs this interest of yours, even if you remain ignorant of it. And so it

36 Cf. Valentini, *Morality and Socially Constructed Norms*, 58.

37 A reviewer worries that this overgeneralizes: it means that I have an interest in controlling, e.g., other people's property. I am unsure about whether this is overgeneralization. But if the reader thinks it is, there are various ways to restrict the relevant interest. We might restrict it by saying we have an interest in controlling only parts of the world that we have a reasonable expectation of controlling, where the reasonableness of an expectation is spelled out epistemically.

does harm you in some respect. By harming you in this respect, it wrongs you. We can treat many cases of putatively harmless wrongdoings with a properly expansive account of people's interests.

3.2. *Pointless Norms*

Let us look at a second challenge. The fair share theory of conventional obligations gives us a reason to obey only norms that have a point, that genuinely help us to discharge one of our collective obligations. I have said this is a virtue of the theory. Yet one might think that there are many norms we should obey that do not have such a point. At dinners in Oxbridge colleges, for example, people must stand while grace is read out before their meal. One might be inclined to view this (and perhaps several other Oxbridge norms) as entirely pointless. Standing while grace is read out, one might think, does not help us to discharge any of our collective obligations. Yet still, if you visit an Oxbridge college, you should stand for grace. The agency respect theory can explain this: the norm of standing for grace expresses the agency of members of the college, and so standing is required to respect this agency. The norm might have no point, but one should respect it nonetheless. So perhaps these cases are evidence against the fair share theory.

I doubt it. The fair share theory can explain why we should obey *seemingly* pointless norms. The simple observation is that such norms are often much less pointless than they appear. Standing during grace is, among other things, a ritual; it is something that all the members of the college do together. Rituals have many beneficial effects.³⁸ Perhaps most importantly, they promote group cohesiveness. They increase the emotional bonds between participants and the likelihood that participants will make sacrifices for the good of the group. We may well have collective obligations to help bring about such group cohesiveness, or at least collective obligations the satisfaction of which is served by such cohesiveness. And so standing at grace may well help us to satisfy our collective obligations. Expressive considerations also furnish a point for many superficially pointless norms. Think about men taking off their hats when they enter a church. This expresses respect for the relevant religion. We plausibly have collective obligations to not disrespect people's religions, so the norm helps us to discharge a collective obligation. There are many ways in which norms that seem pointless on their face have a point on deeper investigation. And so it is perfectly explicable on the fair share theory why one should obey such norms.

Let me address one worry about this line of thought. The line of thought commits me to a degree of liberalism about when we have collective obligations.

38 For an overview, see Xygalatas, *Ritual*.

One might worry that I am too liberal. The worry is that there are beneficial systems of social norms that we do not have a duty to uphold. Suppose, for example, that your neighbors decide to set up a system of public entertainment on local radio. The system has certain norms: that every community member should spend a day running the radio station. But you never agreed to the system being set up. Everyone does benefit from this system of public entertainment, but you (plausibly) do not have a moral obligation to contribute your time to running it.³⁹ The worry is that if we are liberal about what collective obligations people have, then the fact this system benefits people establishes a collective obligation to set it up. And so the fair share theory will implausibly imply that you should help run it. Our response to the problem of pointless norms, the worry goes, overgenerates obligations.

I do not think that is correct. I have said that we might have collective obligations to promote group bonds or to express respect for people's religions. But that does not imply we have collective obligations to entertain our community with a system of public entertainment. That is because, as discussed in section 2 above, we have a collective obligation to fulfil only morally serious goals. Mildly alleviating community members' boredom is not particularly morally serious. Expressing respect for their deeply held beliefs is morally serious. The underlying point is that the line of thought I have advanced does not commit us to the claim that we have collective obligations to bring about every benefit we are capable of bringing about. It commits us only to the claim that we have collective obligations to bring about benefits associated with morally serious goods. So it does not imply that individuals have a duty, in particular, to contribute to public systems of entertainment that their neighbors unilaterally set up.

3.3. *Secrecy*

Let us consider a third challenge. Suppose you are able to violate a norm in complete secrecy. Perhaps, for example, it is a norm in one's society that one pays one's taxes in full. But you can easily get away with underpaying your taxes. You have received some cash payments, and you can refrain from declaring them without telling anyone that you are doing it. Here, we might deny that you actually paying your taxes in full contributes to upholding the norm that people pay their taxes in full. If nobody ever finds out that you did not pay, it seems nobody else will be dissuaded from paying their taxes. And it seems one can continue to criticize others for nonpayment even when one is oneself a tax cheat. But then we might resist the idea that paying your taxes is doing your fair share to uphold the norm. It makes no causal contribution to maintaining the

39 Cf. Nozick, *Anarchy, State, and Utopia*, 93.

norm. So the fair share theory seems to imply incorrectly that if one can secretly violate a norm, one has no moral obligation to conform to it.

In response to this, we need to reiterate what exactly one is required to do to contribute to maintaining a norm. One is not required merely to comply with the norm. Instead, one is required to internalize it. Internalizing a norm means taking it as a rule for your own and others' actions. Internalization is required because it is what most supports the norm—internalization most robustly supports people's enforcement of and conformity to the norm.⁴⁰ We can now diagnose the issue with secretly violating norms. If you have actually internalized a norm, you will not secretly violate it in the way described. If you really take the norm to be a standard for your own behavior, you will not even violate it secretly. So the problem with secretly violating the norm is that it implies you have fallen short of your duty to internalize the norm. In other words, internalizing the norm implies that you comply to it, and if you are obligated to internalize the norm, you are also obligated to comply to it—whether or not anyone will know about your norm violations. Noncompliance means lack of internalization.

3.4. *Discretion*

Let me address a final challenge to the fair share theory. I have suggested that maintaining norms often serves morally serious goals, so one has a duty to do one's fair share to maintain those norm. But suppose that on some occasion, internalizing a norm is not the best way to promote the goal it ultimately serves. Consider tipping at high-end restaurants. Suppose that quite generally, the tipping norm serves the goal of ensuring waitstaff are paid sufficiently for their work. Yet suppose the waiters at high-end places are already well paid. So you would better serve the goal of ensuring workers are sufficiently paid by saving your money and tipping more elsewhere. The concern is that this means that the fair share theory implies incorrectly that you do not have a duty to tip at fancy establishments.⁴¹ More generally, the worry is that the fair share theory gives people more discretion than they have about whether to comply with social norms. We do not generally have a choice about whether to comply with social norms, but the worry is that the fair share theory says that we do.

We can straightforwardly address this version of the discretion worry. The fair share theory does not only say we have collective obligation to serve some ultimate moral goal. As outlined in section 2, it says we have collective

40 For more on the import of internalization, see Kelman, "Compliance, Identification, and Internalization"; and Henrich, *The Secret of Our Success*, 189.

41 I interpret this as the worry in Valentini, *Morality and Socially Constructed Norms*, 65n17.

obligations to take the reliable means to that goal. So we have a collective obligation to maintain social norms that serve the relevant goals. Thus, we have a duty to do our fair share to maintain those norms—not merely to promote the ultimate goal. We cannot do our fair share to maintain a social norm by contravening it. You do not help maintain the tipping norm by not tipping. And so even if not tipping would better serve the goal of ensuring workers are sufficiently paid, you should still tip. If you contravene a norm, you violate your fair share obligations.

Still, there is a version of the discretion worry that is more difficult to address. I have said that we have a duty to contribute our fair share to the maintenance of social norms. But one might do this without internalizing those norms. Suppose I spend all my money on an advertising campaign reminding people to tip and work tirelessly to set up organizations promoting tipping. This might contribute more to the maintenance of the tipping norm than would tipping myself. So one might think that this exhausts my fair share contribution to maintaining the norm. Yet I would do something wrong if I ordered a meal and did not tip. I do not have discretion over how to contribute to maintenance of the norm: at minimum, I have to comply with the norm, no matter how else I contribute to its maintenance.⁴² Is this compatible with the fair share theory of conventional normativity?

I think it is, but to explain it, we have to supplement the fair share theory with some relational egalitarian ideas. The background to these ideas is that some relationships are relationships of superiority and inferiority. Think of the relationship between a Brahmin and a Dalit. Their respective positions in the caste hierarchy put the Brahmin above the Dalit; the hierarchy makes the latter a social inferior. And this wrongs them. We have a claim against being subjected to inferiority. Niko Kolodny is perhaps the foremost theorist of such relationships. He thinks of these relationships in terms of regard.⁴³ You can, he thinks, show someone relatively low regard by esteeming them less than others or by caring about their interests less. Kolodny thinks that relationships of inferiority are partly constituted by differential regard. When one shows one

42 Wellman discusses this issue at length in defending a fair share theory of political authority (“Samaritanism and the Duty to Obey the Law”). Wellman thinks that exercising discretion is valuable, so those who exercise discretion sacrifice less than those who comply with the laws (40–46). I do not think this is a successful response to the worry. Discretion can surely be traded off against other goods. If I spend all my time and money supporting laws but sometimes break them, then I sacrifice more in supporting the laws than do mere law-abiders. I may enjoy a bit more discretion, but the value of this is outweighed by my other sacrifices. *Mutatis mutandis*, the same goes for norms.

43 Kolodny, *The Pecking Order*, 101–16.

person less regard than another, one puts the first person in a relationship of inferiority, thereby wronging them.

For our purposes, the key further thought is that when you encourage other people to obey norms that you yourself do not obey, you elevate yourself above them. That is because such actions show them less regard than one shows oneself. To see this, suppose I always demand that my colleagues pay meticulous attention to their appearance, but I never pay any attention to my own appearance. I criticize colleagues who come to work in casual clothes or without makeup, but I always appear slovenly. Here, I am demanding other people keep to norms that I do not keep to myself. This is a way of putting myself above these other people, of showing myself more regard. And so it subjects them to a relationship of inferiority. This is true even if there is no independent moral reason to comply with these presentational norms: it might be that nobody is antecedently morally obligated to pay meticulous attention to their appearance. Still, if I expect others to dress well while I dress slovenly, I put myself above them. And this is true even if I undertake more of a burden than they do to uphold norms about good self-presentation overall. Perhaps I report my colleague's poor self-presentation to human resources, at considerable cost to my working relationships. Perhaps I even give them money to dress better. Still, by holding them to norms I spurn, I put myself in a position of superiority.⁴⁴

These ideas explain what goes wrong in cases in which someone contributes to maintaining a norm without complying with the norm. That person puts themselves above those they expect to comply with the norm, which violates these people's relational claims. That, then, is our solution to the more difficult discretion problem. Those who promulgate norms without keeping to them put themselves above those who they expect to keep to those norms. Moreover, clearly, if it is correct, the ideas behind the fair share theory do not explain the moral force of social norms in the cases that generate the more difficult discretion problem. That is a weakness of the fair share theory. It is better to explain more rather than less. But I do not think it is a very serious weakness. This more difficult discretion problem hinges on cases in which someone contributes a lot to maintaining a norm but does not comply with norm. Such cases strike me as *recherché*—I suspect they happen rarely in real life. So they are less crucial to explain than more common cases: the core explanatory target of the fair share theory is the most commonplace cases in which social norms have moral force.

44 What if I hold my colleagues to demanding self-presentation norms precisely because I think *they* are better than *me*—and so they should dress more neatly? Perhaps then these relational egalitarian ideas do not apply. But in this case, if I contribute to maintaining the norm in other ways, it strikes me as unclear that I am obligated to myself obey the norm.

Thus if the fair share theory explains the moral force of social norms in more common contexts, the theory hits its explanatory target.

Still, one might wonder whether these relational egalitarian ideas can explain the moral force of social norms in these other contexts.⁴⁵ If so, that would provide a competing explanation for the fair share theory. But I doubt these relational egalitarian ideas have such explanatory potency. The problem is that often norm breakers have little inclination to maintain the norms they break. Reconsider the tipping norm. I might refrain from tipping because I generally oppose tipping. In this case, I do not expect others to tip, nor do I try to get them to conform to the tipping norm. And so there is no sense in which, when I do not tip, I show less regard to other people than I show to myself. Similarly, when Italians come to London, they may have no inclination to enforce the queuing norm. An overly nosy English aunt may have no intention of promoting restrictive English familial norms. Norm breakers are typically dissenters rather than hypocrites. The relational egalitarian thought has little traction in these cases, and so provides a very limited explanation of the moral force of social norms. For a more general explanation, we need the fair share theory.

That addresses what seem to me the most serious challenges to the fair share theory of conventional obligations.

4. CONCLUSION

Why should you queue up at London bus stops, tip 20 percent in New York City, or refrain from criticizing your English nephew's bachelorhood? I have suggested that this is not because we must respect the agency of those who support such social norms. Rather, it is because such norms serve genuine moral goals, and we should do our fair share toward maintaining norms that contribute to these goals. The critical idea behind this view is really a conception of human nature: that we are norm-following creatures. This conception has emerged from the work of anthropologists and other social scientists over the last several decades. It underpins the idea that social norms are causally special: they are the most reliable means of achieving various serious moral goals. This idea drives the fair share theory of conventional obligations. This theory can explain the core cases in which social norms have moral force, and we can address the most serious challenges to it. I believe it is the most plausible account of why we have an obligation to obey social norms.

45 For this kind of view, see Monti, "Are There Any Conventional Obligations?"

I have said little about the wider applications of the fair share theory in this article. But in fact, I think it explains a whole swathe of moral phenomena. We can understand rights quite generally—autonomy rights, property rights, privacy rights, and so on—in terms of the fair share theory. The initial observation behind this understanding is that we have collective obligations that correlate with these rights. We have collective duties, for example, to ensure people have control over their bodies. Fulfilling these collective obligations is very difficult due to the motivational and epistemic issues adduced in section 2. Enter social norms. Setting up or maintaining social norms that constrain people's behavior is the most reliable means of satisfying these collective obligations. So collectively, we have a duty to support such norms, and individually, we have a duty to do our fair share in supporting them. Our duty to respect people's rights is a duty to do our fair share in support of good norms. Of course, much more needs to be said to spell out and defend this conventionalist theory of rights. More, for example, needs to be said about what our moral obligations are in the absence of social norms. But that I leave as a task for another occasion.⁴⁶

Australian Catholic University
adam.lovett@acu.edu.au

REFERENCES

- Chudek, Maciej, and Joseph Henrich. "Culture-Gene Coevolution, Norm-Psychology and the Emergence of Human Prosociality." *Trends in Cognitive Sciences* 15, no. 5 (2011): 218–26.
- Cornell, Nicolas. "Wrongs, Rights, and Third Parties." *Philosophy and Public Affairs* 43, no. 2 (2015): 109–43.
- Dietz, Alex. "What We Together Ought to Do." *Ethics* 126, no. 4 (2016): 955–82.
- Gilbert, Margaret. *A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society*. Oxford University Press, 2006.
- Hart, H.L.A. "Are There Any Natural Rights?" *Philosophical Review* 64, no. 2 (1955): 175–91.
- Henrich, Joseph. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press, 2016.
- Huemer, Michael. *The Problem of Political Authority: An Examination of the*

46 For helpful feedback on this material, I would like to thank an audience at the University of California, Berkeley; an audience at the London School of Economics; two anonymous reviewers for this journal; and especially Laura Valentini.

- Right to Coerce and the Duty to Obey*. Palgrave Macmillan, 2013.
- Kelman, Herbert C. "Compliance, Identification, and Internalization: Three Processes of Attitude Change." *Journal of Conflict Resolution* 2, no. 1 (1958): 51–60.
- Kolodny, Niko. *The Pecking Order: Social Hierarchy as a Philosophical Problem*. Harvard University Press, 2023.
- Marmor, Andrei. *Social Conventions: From Language to Law*. Princeton University Press, 2009.
- Maskivker, Julia. *The Duty to Vote*. Oxford University Press, 2019.
- Melenovsky, C. M. "Promises, Practices, and Reciprocity." *Philosophical Quarterly* 67, no. 266 (2017): 106–26.
- . "The Reasons to Follow Conventional Practices." *Australasian Journal of Philosophy* 102, no. 3 (2024): 710–25.
- Miller, Seumas. *The Moral Foundations of Social Institutions: A Philosophical Study*. Cambridge University Press, 2009.
- Monti, Ezequiel. "Are There Any Conventional Obligations?" *Legal Theory* 29, no. 2 (2023): 90–121.
- Murphy, Liam B. *Moral Demands in Nonideal Theory*. Oxford University Press, 2003.
- Nozick, Robert. *Anarchy, State, and Utopia*. Blackwell, 1974.
- Owens, David. *Bound by Convention: Obligation and Social Rules*. Oxford University Press, 2022.
- Parfit, Derek. *Reasons and Persons*. Oxford University Press, 1984.
- Rakoczy, Hannes, Felix Warneken, and Michael Tomasello. "The Sources of Normativity: Young Children's Awareness of the Normative Structure of Games." *Developmental Psychology* 44, no. 3 (2008): 875–81.
- Rawls, John. *A Theory of Justice*. Harvard University Press, 1971.
- Simmons, A. John. *Moral Principles and Political Obligations*. Princeton University Press, 1979.
- Tosi, Justin. "Rethinking the Principle of Fair Play." *Pacific Philosophical Quarterly* 99, no. 4 (2018): 612–31.
- Valentini, Laura. *Morality and Socially Constructed Norms*. Oxford University Press, 2023.
- . "Respect for Persons and the Moral Force of Socially Constructed Norms." *Noûs* 55, no. 2 (2021): 385–408.
- Wellman, Christopher. "Samaritanism and the Duty to Obey the Law." In *Is There a Duty to Obey the Law?* edited by Christopher Wellman and John Simmons. Cambridge University Press, 2005.
- Xygalatas, Dimitris. *Ritual: How Seemingly Senseless Acts Make Life Worth Living*. Profile Books, 2022.

INTRODUCING DISCORD

Mark Schroeder

IN THIS ARTICLE, I introduce and explain an underappreciated but, as I will argue, pervasive phenomenon that I call *discord*. Discord, I will argue by illustration, helps to explain the source, dynamics, and resilience of many forms of interpersonal conflict. And it is a kind of misunderstanding into which philosophy turns out to offer a particularly privileged form of insight. By better understanding the nature of discord, we can better understand its inevitability, better navigate it, and better appreciate how it can amplify minor conflicts into more significant forms of strife.

1. CONCEPTS

1.1. *Discord*

The phenomenon of discord is simple. Philosophers of action and theorists of responsibility have distinguished between actions for which you are *attributively responsible* and those for which you are not—what is *attributable* to you, for short, and what is not. Discord is my name for what happens when there is a mismatch between how I apply the attributable/nonattributable distinction to you and how you apply it to yourself. That is it. That is the whole phenomenon (up to substituting other people for you and me).

Discord, I will show, is not just a theoretical possibility—it is inevitable. If you and I are in discord, then there is some difference in what we identify as attributable to you. So for us to be in discord, one of us must be wrong. Unfortunately, it is inevitable that each of us will sometimes be wrong about what is attributable to you, because no one—not even you—is infallible about what is attributable to you. This does not quite make discord inevitable, for our mistakes about attributability could be correlated. We could be harmoniously out of tune. But I will argue that not only are our mistakes about attributability not perfectly correlated, but, in fact, we are wired to disagree about attributability. For interpreting what is attributable to someone requires applying interpretive *charity*, and charity is a bias toward the good. So we are bound to disagree about it just as much as we disagree about the good.

But discord is not just inevitable. It is also impactful—I will show that it has consequences. What we interpret as attributable to someone affects how we respond to them. So if you and I are in discord, then I will not respond to you in the ways that you yourself think are appropriate. Worse, unrecognized discord has consequences of its own. If you do not recognize that we are in discord, then when I respond to you in ways that you think are inappropriate, you will infer the wrong things about my motives. And sometimes, as I will show, discord itself persists precisely because it is not recognized.

Fortunately, by giving you the concept of discord in this article, I am equipping you to be able to recognize when you are in discord. So this concept can be therapeutic because it can help you to avoid the bad effects of unrecognized discord and to escape discord that persists only because it is unrecognized. But unfortunately, even after I give you this concept, you will not *always* be able to recognize when you are in discord. Discord can be particularly hard to talk our way out of because the very thing that leads us to disagree about what is attributable to someone—the *charity* with which we apply the concept of attributability—can also give rise to *clumps* of coordinated discord about different topics. And sometimes the topics over which our discord is coordinated include our own attempts to talk our way out of discord.

1.2. *Attributability*

In his classic introduction of the vocabulary of attributability and efforts to distinguish it from closely related concepts in the theory of responsibility, Gary Watson says that you are attributively responsible for some action when it in some sense expresses your true self, as you determine it.¹ Watson follows John Dewey in adding the ‘as you determine it’ clause. This clause fairly accurately describes a wide class of philosophical theories of attributability, from Dewey’s own, to Harry Frankfurt’s, to Michael Bratman’s, Christine Korsgaard’s, and David Shoemaker’s.² All of these theorists say not just that some actions in some sense express your true self but that you in effect get some *say* in what your true self is—either by accepting some things about yourself or by identifying with them, them resonating with you, or fitting into more comprehensive planning structures, or the like. But my own view is that all of these views misidentify the kind of power that we each have over our true selves.³ So I

1 Watson, “Two Faces of Responsibility.”

2 See Dewey, *Outline of a Critical Theory of Ethics*; Frankfurt, “Freedom of the Will and the Concept of a Person”; Bratman, *Structures of Agency*; Korsgaard, *Self-Constitution*; and Shoemaker, *Responsibility from the Margins*.

3 See Schroeder, *When Things Get Personal*, especially ch. 5. Even if you do not share my view, it is better to work with a less contentious concept so long as it is easy to do so.

propose to leave this out. Let us say, then, that the actions that are attributable to you are those that express your true self, and let us leave as a separate question what makes something part of your true self, including whether you get any say in what determines this, as well as how seriously to take the metaphor of a “true self.”

This definition is only as helpful as the metaphor of the true self, so examples will be helpful. In a classic pair of cases, Frankfurt distinguishes between the willing and unwilling addicts.⁴ Both have a powerful addictive desire to take their drug—so powerful that it is inevitable that they will succumb before the end of the day. But the willing addict rises from bed eager to get their first hit and structures their day around it, whereas the unwilling addict awakens in the hope that today is day one of being clean and spends most of their day taking all of the right steps to make this happen—destroying their stash, throwing away their needles, deleting their dealer’s contact info, and logging into an online addiction recovery support group. (Of course, eventually they lose their nerve—he did warn us up front that it was inevitable.) Frankfurt thinks that we can see the difference between these two characters, which he describes as a difference in which acted *freely*. Watson identifies the concept of freedom in which Frankfurt was interested as a paradigm of trying to understand attributability. The willing addict—but not the unwilling addict—is attributively responsible for taking the drug.

Here is another example that I like more.⁵ When we get together to discuss this paper, you ask me a question, and I snap harshly at you, “No!” At first, you might get angry at me for being rude. Or, depending on your personality, you might instead get anxious that you have made some mistake that I am annoyed at—or even, adopting this hypothesis, become embarrassed about it. But instead, it might occur to you that we are having this conversation in midafternoon, and I have not had a chance to grab lunch. Perhaps I am merely hangry, and rather than getting angry at me or embarrassed, you should just pass over it and steer our conversation toward where we can find a snack.

Mars, Inc. has founded a successful international advertising campaign supporting over \$450 million in annual sales on their bet that all of us recognize that we are not always fully ourselves, and hunger is a familiar—and relatively easy-to-manage—way in which we can fail to be fully ourselves.⁶ Their ads feature crabby, belligerent, and difficult people on rampages until someone hands them a Snickers bar, at which point, no longer hangry, they suddenly turn back

4 Frankfurt, “Freedom of the Will and the Concept of a Person.”

5 See Schroeder, “Tipping Points.”

6 Beadle, “America’s Top 10 Best-Selling Candy Bars of the Year.”

into themselves (depicted cinematically by the substitution of a different actor). These ads work because we all recognize the idea that not everything that we do reflects (“expresses”) who we really (“truly”) are. They directly evoke the philosophers’ metaphor of the “true self.” But they also show that this distinction is not just one that philosophers make after reading some Dewey or carefully attending to patterns in pairs of cases. It is a distinction that ordinary people make—ordinary enough for the Mars corporation to bet big on selling them candy bars in this way.

1.3. Participant Responses

So why do ordinary people make this distinction? I think that the answer is simple. Attributability is, I suggest, the “in” to what we can call, following Peter Strawson, *participant responses*. In “Freedom and Resentment,” Strawson noted that there is a large variety of ways that we relate to persons but not to other kinds of thing—ways that we relate to *whos* rather than *whats*. When you are angry, for example, I can ask *what* you are angry *about* but *who* you are angry *at*. Anger is, in Strawson’s terms, a participant attitude because it answers to a *who* rather than a *what*.⁷

All the participant responses that Strawson discussed are attitudes. So he does not distinguish between participant attitudes and other kinds of participant responses. He also endorsed a strong thesis about what the participant attitudes have in common. He said that they are all *reactions* to someone else’s attitudes. So he conflated both of these distinctions, referring only to what he called the *participant reactive attitudes*.⁸ But we should make both of these distinctions. Even if it turns out that participant attitudes are all reactive, we should distinguish that as a substantive further thesis that requires additional support. And there are many clear examples of participant responses that are not attitudes at all—any verb that answers to “*who?*” rather than “*what?*”

Take complaining, for example. I can ask what you are complaining about but only *who* you are complaining *to*.⁹ Yet complaining is not an attitude—it is a speech act. Or take the example of listening. True, there is a general kind of listening that we can do with music, ocean waves, or the creaking of the stairs. But that is not the kind of listening that we seek from loved ones or therapists. We want from them a distinctive kind of listening that we do to persons but

7 Strawson, “Freedom and Resentment.” Compare Langton, “Duty and Desolation”; and Schroeder, “Persons as Things.”

8 Compare Holton, “Deciding to Trust, Coming to Believe,” who first coined the term ‘participant stance’ to pick out the first of these Strawsonian ideas.

9 Compare Bosco, *The Triangle of Innocence*.

not to things. Listening is also not just an attitude. So the class of participant responses is potentially quite diverse.

Strawson notes that it is possible to detach from the kind of perspective that we need to occupy in order to have participant responses. You can step back and observe someone from a more objective, clinical perspective, from which you will not get angry at them or listen to them. You can think of them as a what rather than as a who. And he also allows that you can exclude some things from your participant responses to someone. For example, while listening to what they say, you can take the objective perspective toward their tone of voice—thinking of it as just a result of being hangry. When you do so, you are thinking of what they say as reflecting who they are, but their tone of voice merely as reflecting what they are.

Strawson makes it sound like these are more and less extreme versions of the same thing—that when you exclude aspects of what someone does from the participant perspective, this amounts to a kind of *restriction* or *limitation* on the participant perspective to seeing someone as a what rather than a who (though to a more limited extent). But I do not think that that is quite right.¹⁰ We are all of us embodied in imperfect ways. We are subject to hunger and hormones. It is not a limitation on seeing you for who you are to recognize and appreciate the limits of your embodiment. I can see you *better* as who you are if I am prepared to recognize the limits imposed by your embodiment. So omitting some things from my participant responses to you is not necessarily a way of seeing you as a what rather than as a who. It is sometimes required in order to see you for who you are.

The connection between attributability and participant responses brings us back to our earlier metaphor that what you are attributively responsible for is a matter of what expresses your true self. Dropping the redundant word ‘true’, what is attributable to you is what reflects *who* you are. Actions for which you lack attributive responsibility, in contrast, reflect *what* you are but not *who* you are—they come from your embodiment, as someone who has an addiction or has missed lunch. So just as we do not have participant responses to rocks or rainbows, we do not have them concerning actions that we do not attribute to you. The actions that we respond to—that we listen to, get angry about, are proud of, or respect, among many other participant responses—are limited to those that we interpret as attributable to you.¹¹

10 Compare Schroeder, “Persons as Things.”

11 And this in turn explains why accountability entails attributability, as theorists of responsibility often allow but often leave unexplained. To be accountable is to be fittingly called to account. But calling someone to account for something is a participant response to

So ordinary people need this distinction, I suggest, for the same reasons that they need to grasp the distinction between whos and whats—it shapes our ordinary interpersonal relationships in pervasive ways by enabling modes of response that we do not have to mere things. And importantly, these modes of response are diverse. They include not just attitudes like anger and resentment, which have received so much attention in the theory of responsibility, but also ones like pride, gratitude, and appreciation, as well as other sorts of actions like *listening to* what someone says and *honoring* her requests.

2. CONSEQUENCES

2.1. Error

From the fact that what we attribute to someone shapes how we respond to them, it follows that mistakes in attributability interpretation are not idle. If we make mistakes about attributability, then that has consequences for how we relate to someone. And those consequences can shape our relationship with them in unfortunate ways.

Suppose, for example, that I snap at you simply because I am hangry. My snap does not mean anything—there is no broader import to it or anything that it reveals about how I really feel about you. I am just crabby because it has been a few hours since I have eaten, and you get the brunt of it. The success of the Snickers advertising campaign turns on our familiarity with the idea that in at least some cases like this, my snap is not attributable to me. So let us suppose that ours is such a case. If you correctly identify this, then my hanger will cause us only minimal trouble. Overlooking it, you can pass me a Snickers bar or remind me that we should break for lunch, and we can move past it without incurring any lasting effects on our relationship.

But if you mistakenly think that my snap is attributable to me, then things will not go so smoothly. Now my snap is eligible for participant responses. You may get angry at me for my rudeness or embarrassed about what mistake you may have made that I am responding to. If in fact, however, all that is going on is that I am a bit hungry, then this imperfection in my embodiment is getting in the way of our relationship and of our understanding one another. You are getting angry or embarrassed about something that is not worth getting angry or embarrassed about. You are *overprojecting* attributability.

them. So you can have this response only to what you interpret as attributable to them. For further development of this point, see Schroeder, *When Things Get Personal*.

The ability to distinguish between what is attributable to someone and what is not is so important because it allows us to avoid some of this kind of mistake. If we have the concept of attributability, we are not doomed to overreact to everything that one another does. But having this concept also creates the risk of a new kind of mistake. It creates the possibility of *underprojecting* attributability, interpreting some action as not attributable to someone when it really is.

Suppose, for example, that something that you have been doing has been bothering me for months. But every time I try to tell you about it, you are simply so charming and I am simply so afraid of conflict that I am unable to go through with it, and so unwittingly, you continue to do this thing that bugs me (mispronouncing my name, for example). Today, I have missed both breakfast and lunch, and so due to being hangry, I am finally crabby enough to overcome my timidity about conflict and tell you what I really think, even though you are so disarmingly charming. But unfortunately for me, as you have gotten to know me, you have learned to recognize the signs that I am hangry. So when I tell you what I really think, you just pass me a Snickers bar.

Something goes wrong in this case, but it is different from what goes wrong when you overproject attributability. Instead of responding to *too many* things about me, now you are responding to *too few*. But it follows from the fact that we exclude some things from participant responses (as I have argued, by interpreting them as not attributable—but this label is not essential for the structure of the problem) that we can at least potentially make the mistake of excluding too many things—and hence make the mistake of responding to too few.

So far, I have not said anything about what the attributable/nonattributable distinction is really about—what it really tracks. We have just identified it with the metaphor of expressing your true self, applied it to some paradigmatic examples, and established its role in connection to participant responses. But *whatever* the nature of this distinction turns out to be—whatever it is really tracking—it will be something that it is possible to be wrong about. Mistakes about attributability are inevitable. And as I have shown, these mistakes have consequences for how our relationships go.

Nothing about the inevitability of these kinds of mistakes tells us how common we should expect them to be. Some things are easier to know about—and hence easier to avoid mistakes about. If attributability is one of the things that it is easier to know about, then we should expect these kinds of errors to be less common. But if it turns out to be one of the things that it is harder to know about, then it makes sense to expect these kinds of errors to be more common. When philosophers theorize about attributability, they normally do so directly. They do not worry about what kinds of mistakes people are prone to make about it. But I suggest that we can get leverage on understanding what

kind of thing attributability is from the other direction. We can observe how often people make mistakes about it.¹²

My conjecture is that now that we have seen what the consequences of mistakes about attributability look like, you will agree with me that the circumstantial evidence suggests that these kinds of mistakes are quite common indeed. Respecting, acknowledging, listening to, and being proud of are all participant responses. So whenever someone does not respect, acknowledge, or listen to you or fails to be proud of you when you expect it, you are experiencing the symptoms that we would expect if they were making the mistake of underprojecting attributability to you. Women who are dismissed as hysterical, Black men who are dismissed as angry, and teenagers who are dismissed as hormonal all experience what we would expect it to look like if others underproject attributability in their interpretations of them.¹³

2.2. *Discord*

Actually, this is not exactly right. When you feel like someone is not listening to you, it *looks to you* like you are experiencing the expected symptoms of them underprojecting attributability onto you. But this is also what it would look like to you if they were actually interpreting you correctly and *you* were the one *overprojecting* attributability onto yourself. If you have ever apologized to someone and admitted that they were right to pass you a Snickers bar or wait to rediscuss an issue after you had sobered up, then you know that when you are in the heat of being affected by hanger or alcohol, part of the experience can be precisely that of *not* being inordinately affected in these ways. And things only become clearer once you are no longer hangry or drunk. So you know that from the inside, it can look like the other person is overlooking or dismissing you even if they are not.

So in the first instance, our experiences of the symptoms of another person underprojecting attributability onto us are better evidence that our interpretations of what is attributable to us do not match the person's interpretation of us than they are that the other person is making a mistake. This mismatch is what I call discord. The fact that discord involves *mismatch* makes it a better concept to use, I believe, in order to understand interpersonal conflict. If you make a mistake about what is attributable to someone, and they make the very same mistake about themselves, then in a way, both of you misunderstand

12 See Schroeder, *When Things Get Personal*, especially chs. 4–9.

13 This does not mean that this is *all* that is going on in these kinds of distinctively gendered and racialized experiences—that is most certainly not the case. And it does not *entail* that discord plays any role in these experiences. But it is striking evidence that errors of attributability play *some* important role in *many* of these experiences.

something important about them and about your relationship. But because you misunderstand it in the very same way, this is not going to cause you trouble in getting along. This is like being out of proper tune but in harmony with one another. Disharmony arises not when someone gets out of tune but when not everyone goes out of tune in the same way, together.

Because discord is a mismatch in interpretation, it requires error. But the concept of discord is neutral about who is making the error. This makes the concept of discord an especially useful tool, I think, for perspective-taking in conflict. If two people are in discord, then it looks to each of them *as if* the other person is engaging in inappropriate participant responses. Applying the concept of error in the same circumstances focuses our attention on the wrong thing, in order to understand what things look like, from their perspective. It focuses our attention on their being wrong. And so if this is our customary way of thinking through these cases, it makes it harder for us to appreciate that we may be the ones who are wrong.

Because discord happens only when there is error, it could be that the explanation of discord always consists in the explanation of particular errors. You and I might be in discord, for example, because I make a particular mistake. If that were right, then thinking about discord would always lead us back to thinking about error, as soon as we started wondering why we are in discord. But later in this article I will argue that this is not true: sometimes discord can be explained *directly*, without explaining either person's error. This is because, I will argue, attributability interpretation is *value laden*, and so disagreements in values will engender disagreements in attributability interpretation. So we can sometimes have a pretty good understanding of where discord comes from without yet getting to the bottom of the question of who is the one who is in error.

2.3. *Unrecognized Discord*

When you and I are in discord, the way that I respond to things looks, from your perspective, to be inappropriate. When I get angry, for example, you do not identify the thing that I am angry at as something to which anger is an apt response. Or you expect me to respect your expression of your wishes, but I do not. This, I have argued, can have problematic consequences for our relationship. But unrecognized discord is worse.

When two people are in unrecognized discord, their attributability interpretations do not match, but they do not realize that their attributability interpretations do not match. Even though you are in general aware that other people do not believe all of the same things as you do, when something seems especially obvious, it can be especially surprising that it looks a different way to someone else. When the infamous photo of "the dress" took over the internet in February

2015, for example, what baffled people who saw the dress as gold and white was how anyone could possibly (and seriously) see it as black and blue.¹⁴ When you look at a photo of a gold and white dress, it does not normally occur to you to wonder whether someone else is taking for granted that it is black and blue.

While I was writing this article, for example, I ran an errand to pick up some gold-colored gift bags for my daughter to distribute holiday gifts to her friends, and I handed them to her in the car while we were driving home after dark that evening. She was quite upset at me for buying her black bags for holiday gifts and did not believe that they were in fact gold until we pulled over and had more favorable lighting conditions. Similarly, if an attributability interpretation seems very obvious to you, it might completely fail to occur to you or seem creditable as a serious possibility that the other person sees things another way.

But unrecognized discord creates illusions of ill will. If I can see that you are angry at me, but I do not see the way that I snapped at you as a legitimate object of anger because I had obviously just missed my lunch, then I will have to look for another answer to what you are upset at—and I will arrive at the wrong answer. If I can see that you are not listening to what I say but rather just passing me a Snickers bar in the hopes that I will stop saying it and we can move back to another subject, then I will infer that you must not *care* about what I am saying—or at least, do not care enough. In general, if I have a different space of interpretive possibilities of what you are responding to about me and how, then when I try to understand what beliefs and motives are leading you to respond in these ways, my mistaken understanding of the totality of your beliefs is going to lead me to a mistaken understanding of your motives.

So one way that unrecognized discord creates problems is by leading us to misidentify someone's quality of will. But unrecognized discord can also be worse in a different way. And that is that discord can sometimes persist precisely *because* it is unrecognized. Whether this is so depends on each person's *basis* for applying the attributable/nonattributable distinction in the way that they do. So far, I have not said anything about how we apply this distinction, and in particular, I have not said enough in order to be able to illustrate how the failure to recognize discord can reinforce the underlying attributability interpretations that constitute that discord. So this is just a promissory note. But I want to note it here as a special further deleterious consequence of unrecognized discord.

Fortunately, not all discord is unrecognized. In particular, now that you have the concept of discord, you may recognize it. But even before you had a name for the concept of discord, you may sometimes have appreciated that someone else was interpreting your situation very differently than you were. You may

14 See Resnick, "The Internet Peaked with 'The Dress'."

have used less specific vocabulary to grasp at this—for example, you may have said that they had a “different narrative.”¹⁵ Or you may have had a narrower concept identifying special cases of discord, without having a general concept that encompasses everything that I count as discord. For example, you may have recognized or even had a word for the experience of being a woman making a point in a meeting that is not taken up or acknowledged until it is made again by a male colleague; but you may not have considered it a single concept that includes both this and the experience of a child who is frustrated that his parents are not proud of him for something that they take for granted—even though discord can encompass both of these as special cases. So some discord can be recognized, even if it is not conceptualized specifically in the way that I have.

But much discord is unrecognized. And even once you have the concept of discord, it can remain unrecognized. Discord creates problems, but unrecognized discord is worse.

3. WRINKLES

3.1 *If Only Things Were Simple*

So far I have explained what discord is and why it has certain kinds of characteristic earmarks—a mismatch between the participant responses that people exhibit and those that we expect of them, which can often consist in their not listening to us, not respecting us, not being proud of us, or the like. And I have observed how unrecognized discord can create illusions of ill will. This is the impactfulness of discord. I have also shown that error about attributability is inevitable, so that discord must also be inevitable unless we can coordinate our errors with one another. And I have pointed toward circumstantial evidence that both error and discord are common.

You might think, however, despite the fact that its characteristic symptoms are common, that discord should not itself be quite so common. You might think that it could be easily avoided or at least moderated by acquiring a better understanding of psychology or neuroscience. One way that you might think that we could avoid discord is to always accept each person’s interpretation

15 The concept of discord does not compete with the idea that you and someone else accept “different narratives” about what has happened as a potential explanation of why you each have different emotional responses to what has happened between you. Rather, it tells us what makes differences in narratives relevant and how. As Lindemann emphasizes, narratives work by foregrounding and backgrounding information (*Damaged Identities, Narrative Repair*). They select some events as important and significant. This is exactly what attributability interpretation does. Compare Schroeder, “Narrative and Personal Identity.”

of themselves. It is a familiar idea in popular culture that it is wrong not to take someone's self-interpretation at face value, after all, and on some interpretations of that idea, it might imply that if you know what someone attributes to themselves, you should believe that interpretation, or at least not disagree with it. You might even think that on many of the theories of attributability developed in the philosophical literature, something like this ought to be true because many of those theories take seriously and develop in different ways Dewey's idea, emphasized by Watson, that attributable actions are those that express yourself *as you determine it*.¹⁶ So if what is attributable to you is a matter of what you determine, then who better to know what is attributable to you than you?

But this idea, I think, cannot be right.¹⁷ Each of us, I conjecture, can identify times in our lives when we have realized that we were ourselves wrong about what was attributable to us. At those moments, we did not think that we were merely hungry or hormonal—the issues at stake felt really important to us. It is only in retrospect that we look back and realize that that is precisely how the hanger or hormones got their grip on us—by making those issues feel so important at that moment. More generally, the fact that we interpret ourselves in different ways at different times means that we cannot always be right. And all the proponents of familiar views of attributability according to which it is in some sense self-determined can allow this. Whatever sense in which attributability counts as self-determined, even on these views, is not one that makes it at all times transparent to you what you have self-determined.

Whether or not attributability is self-determined, you might think that it is a matter of an action's having the right sort of cause. It must spring from, say, desire rather than impulse, values rather than desires, or complex planning structures that integrate agency over time rather than one-off plans.¹⁸ Nearly all philosophical accounts of attributability have this structure. What they disagree about is which cause is the special one that makes actions attributable and whether this cause has first-order unity or might instead be unified only at some more abstract level, such as being a cause that you “identify” with or accept or is relatively enough enduring within your psychology to count as “character.”¹⁹ If attributability is a matter of having the right cause, then we

16 Similarly, many philosophical accounts of identities or of personal identity make them in an important sense self-constructed. Compare, for example, Schechtman, *The Constitution of Selves*.

17 Compare Schroeder, *When Things Get Personal*, especially ch. 5.

18 See, for example, respectively, Arpaly and Schroeder, *In Praise of Desire*; Watson, “Free Agency”; or Bratman, *Structures of Agency*.

19 Frankfurt, *The Importance of What We Care About*; and Shoemaker, *Responsibility from the Margins*.

should expect that mistakes about attributability will be more common the less that we understand about human psychology and, correspondingly, less common the more that we come to learn about psychology.

But I conjecture that this is not the case. In fact, I think it is far from being the case. The more we learn about human psychology, the more possible diagnoses we have available to us for understanding the causes of one another's actions. We can form hypotheses about attachment styles, neuropathies, anxiety, depression, and personality disorders. We can keep track of whether or not someone is "off their meds" and about how tired they are. We have specific new concepts like that of being hangry that make it easier for us to identify new kinds of behavior to overlook and manage. All of these tools make it easier and easier for us to diagnose when something is *not* attributable to someone, because they provide us with a longer and more fully articulated list of alternative interpretive possibilities for what it is attributable to instead.

But I conjecture that we do not find less discord now than ten, twenty, or two hundred years ago. On the contrary, experiences of being overlooked, ignored, diminished, objectified, unseen, and unlistened to figure especially prominently in all manner of contemporary literature and are heavily theorized by feminist and intersectional theorists. The problem is that these tools for better psychological understanding make it easier to identify when things are *not* attributable, only at the cost of making it easier to make the mistake of thinking that something is not attributable when it really is. And everyone who has experienced any of these diagnoses firsthand knows that it is not so easy to do so. Just as hunger might sometimes lead me to snap in irrelevant ways but can also sometimes be the very thing that allows me to get over my timidity and your charm and follow through to tell you what I really think, depression can have some effects that are not attributable to you while also having others that are. And the same thing goes for everything else that we might put on our list of helpful diagnoses.

I conclude, tentatively, that things are probably not so simple as this. If the concept of attributability tracked a purely causal, psychological distinction, then we should get better at applying it the more that we learn about psychology.²⁰ And this seems to me to be very far from obviously what we actually do observe. But we can also find direct evidence that attributability does not work quite like this—or at least that we do not think about it in this way.

3.2. *Attributability and Charitability*

Suppose (in a distant counterfactual possibility—she insists that I add just to clarify) that when my wife comes home from work and compliments my

20 See especially Schroeder, *When Things Get Personal*, chs. 6–7.

gardening, I start speculating about what lucky thing must have happened to her earlier in the day that put her in the mood to pay me a compliment.²¹ This would be icky. Something is wrong with our marriage if my response to compliments is not gratitude or pride but speculation about where they came from. Still, the more we know about psychology, the better we can appreciate the truth that people are more likely to compliment one another when they are in good moods and more likely to be in good moods when something favorable has happened to them. The ickiness of my responding to my wife's compliment in this way therefore has nothing directly to do with how likely it is to be true.

In contrast, if my wife instead comes home and (in an even more distant possibility—she insists that I add) complains about my gardening, there is nothing icky about my holding in reserve the hypothesis that she has merely had a bad day. An important part of being able to get along with someone involves understanding their embodiment and hence being prepared to recognize that not everything they do or say reflects on them or on your relationship with them.

But of course the causal connection between complaining and having had a bad day—mediated by mood—is precisely the analogue of the causal connection between complimenting and having had a good day. The contrast between the right way for me to respond interpretively to these cases does not come from a difference in their causal structure. It comes from the difference between compliments and complaints. The difference between these two cases is, I conjecture, an evaluative one. And their contrast reveals that attributability interpretation should be value laden. It needs to be biased towards the good.

I do not mean to say that compliments are always good and complaints are always bad. When complaints add to our understanding of what is genuinely important to one another, they can be overall quite good, even though they are unpleasant to process. But there is also something icky about speculating about what bad event earlier in the day led to a constructive and instructive complaint like this. So I think that the contrast in ickiness is a contrast that tracks an evaluative difference between the cases. Insofar as you agree with my ickiness judgments, you should agree that attributability interpretation—at least between spouses—*ought* to be biased toward the good. Other things being equal, we should lean towards attributing good things and away from attributing bad things.²²

21 This case comes from Schroeder, "Persons as Things."

22 I argue for this claim at greater length in Schroeder, "Persons as Things," "Attributive Silencing," "Tipping Points," *When Things Get Personal*, and *Interpretive Objects*. Here I follow the argument of "Persons as Things." See also Christine Korsgaard's argument (in "Creating the Kingdom of Ends") that responsibility judgments are practical.

But this is not just something that we should do. Even more importantly for my purposes here, it is something that we *do* do. It turns out that when psychologists and empirically minded philosophers have set themselves the task of examining how people actually make judgments about attributability and the related concept of a “true self” that attributable actions are said to “express,” they have found systematically that people do in fact make true-self judgments in ways that are evaluatively biased towards the good. George Newman, Paul Bloom, and Joshua Knobe use several experiments to draw out the way in which people’s true-self judgments are biased towards what they believe to be good; and Newman, Julian de Freitas, and Knobe argue that the value asymmetry in true-self judgments underpins and explains many other value asymmetries that experimental philosophers have uncovered in applying many closely related concepts.²³

So attributability interpretation is and should be charitable. It is and should be biased toward the good. I say that it not only is but *should be*. But all that I need in what follows is that we *do*, as a matter of empirical fact, use a charitable bias toward the good in determining what is attributable to someone. This fact turns out to make sense of many interesting features of the kinds of mistakes that we make in interpreting ourselves and one another—and of many complex and interesting features of discord more generally.

3.3. *Charitability of Attributability Explains Systematic Patterns in Errors*

We observed earlier that underprojecting attributability has the consequence that we can fail to have appropriate participant responses to things that merit it. We can fail to be proud of things that deserve it, to honor or respect people’s wishes, to admire their accomplishments, or to listen, because all of these things are among the many forms of participant response. We treat someone as a little bit less like a who than is called for and a little bit more like a what. But it turns out that this kind of mistake is unevenly distributed.

People are whos. Things—objects—are whats. So treating someone a little bit less like a who and more like a what is treating them a little bit less like a person and a little bit more like a thing—more like an object. Another word

23 Newman et al., “Value Judgments and the True Self”; and Newman et al., “Beliefs About the True Self Explain Asymmetries Based on Moral Judgment.” Some of the asymmetries covered by Newman, de Freitas, and Knobe include asymmetries in what subjects count someone as *valuing* (Knobe and Roedder, “The Ordinary Concept of Valuing”), in who they count as being *happy* (Philips et al., “The Ordinary Concept of Happiness”; and Philips et al., “The Good in Happiness”), in the conditions under which they are counted as experiencing *weakness of will* (May and Holton, “What in the World Is Weakness of Will?”), and in who counts as *blameworthy* or *praiseworthy* (for example Pizarro et al., “Asymmetries in Judgments of Moral Blame and Praise”).

for treating someone a little bit more like a thing and less like a person than they really are is *objectification*. But theorists who study objectification do not in general find that objectification is equally well distributed across different people. Instead, they pay attention in particular to ways in which women are more often objectified than men.

Listening is a particularly important form of participant response. It is through listening to one another that we are able to cooperate and live together. Of course, there are kinds of listening that you can do to instrumental music or to the sound of waves lapping on the shore. But when you tell someone how you are feeling, this is not the kind of listening that you are looking for—you want them to listen to *what you are saying*, not just to how it sounds when you say it. This kind of listening is a participant response. But if someone is telling you things and you are not listening to them because you interpret it as noise rather than as signal, then there is a very natural sense in which they are *silenced* to you. This is what Mary Kate McGowan calls *true-self* silencing.²⁴ Like objectification, we have substantial evidence that the experience of silencing is not equally distributed across people. It has been feminist theorists who have called our attention to silencing and done the most to theorize about it because the experience of feeling silenced is particularly prominent for women in particular.

Giving credit is another form of participant response. When someone does something, we can be grateful for it and thank them; we can admire it and praise them; or we can acknowledge that it came from them (perhaps in our bibliography). Famously, the giving of credit is also not equally well distributed across different people. Who gets credit and what they get it for are famously infected in interesting and complex ways by gender, social status, and economic class, among other things.

Here is an utterly simple conjecture about what explains much of these differences in how objectification, silencing, and the giving of credit are distributed across race, gender, and class, among other social distinctions. It is that this is a consequence of the fact that social values are distorted in ways that correspond to race, gender, and class. I describe this as a conjecture, but notice that it requires making no new assumptions. We all know that social values

24 McGowan, *Just Words*. This notion makes errors of underprojecting attributability a promising way of accounting for some experiences of being silenced that draws on tools from the philosophy of action rather than the philosophy of language (as Langton, “Speech Acts and Unspeakable Acts”; Langton and Hornsby, “Free Speech and Illocution”; and Hesni, “Illocutionary Frustration” do, among many others) or epistemology (as in Dotson, “Tracking Epistemic Violence”). True-self silencing is not so much a competitor for these other tools for understanding silencing so much as a closely related phenomenon that can overlap with other forms of silencing or encompass cases that they fit less well. Compare Schroeder, “Attributive Silencing.”

are distorted in ways that correspond to race, gender, and class. And we have already seen that what people attribute to someone is affected by their values. And finally, we established earlier that attributability is the “in” to participant responses—what we have participant responses to is limited by what we attribute to someone. So we already have substantial evidence for each of these three assumptions. The fact that when put together they predict the utterly banal observation that objectification, silencing, and the giving of credit are likely to be distributed in ways that are affected by race, gender, and class should further increase our confidence in each.

4. MORALS

4.1. *Symmetric Explanations of Discord*

I set out at the beginning of this article to introduce you to the concept of discord. I showed that discord is easy to define from the philosopher’s concept of attributability and that since we all apply the concept of attributability in our everyday relations to one another, it matters whether we are in discord or not. I also used the general connection between attributability interpretation and participant responses in order to explain *why* mistakes about attributability such as those that arise in discord can create trouble for interpersonal relationships. And I explained why unrecognized discord can be especially pernicious. Along the way, I hope that I have illustrated or at least alluded to enough applications of this framework in order for you to see why mistakes in attributability interpretation are particularly important for us to think about and understand—and to begin to anticipate how having the concept of discord might be therapeutic.

The most controversial thing that I have said so far, I think, is that attributability interpretation is properly value laden. I claimed that the right way to interpret what is attributable to someone requires applying a kind of interpretive *charity*, which means that it requires exercising your own values. But even if I am not right about this, we at least have substantial evidence that this is what people in fact do—both direct empirical evidence from the laboratory setting and also indirect evidence in the form of the explanation that it offers of the systematic maldistribution of things like objectification, silencing, and the giving of credit, which are not otherwise well explained by the hypothesis that we are merely applying a simple scientific distinction to one another.

But the reason why I wanted to get to this more controversial claim about attributability interpretation and not merely to settle for introducing the concept of discord in general—in a way that is independent from assumptions about

how we actually interpret what is attributable to one another—is for its theoretical fruits. We have seen one such fruit: it gives us new potential insight into the sources and mechanisms of the systematic maldistribution of objectification, silencing, and the giving of credit, among many others. To fully realize this insight, of course, we need to walk more carefully through *how* attributability interpretation is value laden and to rely more specifically on assumptions about *how* social values are distorted. That is work for another occasion. But I want to close this article with two more important upshots of the fact that attributability interpretation is value laden. The first, in this section, concerns the value of the concept of discord for *perspective taking*. And the second, in the sections that remain, concerns the way that discord can *clump* around related topics.

Earlier, I emphasized that the concept of discord contrasts favorably with the simple concept of error in attributability interpretation in offering a better tool for perspective taking. If you are in a conflict with someone, recognizing that you are in discord offers you a helpful window into how things look *to them*. Because it is a symmetric concept, it focuses our attention on what is symmetric about the situation and hence makes it easier to appreciate that from the inside, the other person could be equally frustrated or mystified about you. And because the concept of discord is neutral about whose interpretation is in error, if we start with discord, there are equally natural paths into wondering whether we are the ones who are mistaken, as wondering why the other person is mistaken.

Of course, if the only ways that we have of understanding that or how we could have gotten into discord with someone start by understanding that or how one of us got attributability interpretation wrong and then abstracting away from who it was, then framing what is going on between us in terms of discord is not particularly helpful for deciding which of us is mistaken. In this context, it is particularly valuable to appreciate how attributability interpretation is value laden.

Value disagreements, we know, are extremely common. Even among people with deeply shared values, there are lots of evaluative questions still to disagree about. The pervasiveness of value disagreements means that since attributability interpretation is value laden, discord is also going to be pervasive. We are bound to find ourselves at least sometimes in discord, because there are bound to be cases in which our value disagreements rise to the fore in our attributability interpretations, even if we agree about very many values. Recognizing the way in which our attributability interpretations are bound to be informed by our values therefore offers a particularly neutral way of recognizing that we may now be in discord—a way that is not mediated by identifying any particular mistake that one of us made. It is a way of thinking about the source of discord that can help to open us up to think about it symmetrically—or even

to acknowledge that the other person sees the situation correctly, and we are the ones who were in error.

4.2. *Unrecognized Discord, Redux*

Earlier, I argued that the effects of discord are often worse when they are unrecognized. The main reason for this is that when you do not recognize that you are in discord with someone, you interpret them as responding to the same things that you see as apt to be responded to. And so if they are in fact responding to something else, then this gives you a misleading impression of their motives and priorities. It creates, as I put it, illusions of ill will. But I also claimed (so far without argument) that sometimes discord persists precisely *because* it is unrecognized. I did not complete the argument for that claim earlier because we did not know enough yet about how people decide what to attribute to someone. But now that I have argued that we use *charity* to interpret what is attributable to someone, we have the necessary piece to complete this argument.

I do not mean to suggest that *all* discord persists only because we are unaware of it. Far from it. The very fact that attributability interpretation is value laden suggests that all it should take for discord to persist is for two people to have different values. Then no matter how much evidence they each acquire, the difference in their values, mediated by the application of the principle of charity, will lead them in different directions. For example, realizing that your father is not proud of you because he does not see your accomplishments as really belonging to you is not a way of becoming convinced that he is right about that. It might help you to appreciate that he really does care about you even though he is not proud of you for these particular things. But you can recognize that the two of you disagree about what is attributable to you while remaining confident that you are the one who is right.²⁵

The reason why failing to recognize discord can help it to persist is simple. It is that discord persists because competing attributability interpretations persist. Competing attributability interpretations persist, when they do, because each of the interpretations persist. And because attributability interpretation relies on charity, an attributability interpretation can persist because it continues to be charitable. But what it is charitable to attribute to someone can depend on what you think that they are doing. And the space of hypotheses about what they might be doing is shaped by the space of things that you think they could

25 If some form of conciliationism in the philosophy of peer disagreement is correct (compare Elga, "Reflection and Disagreement"), then it could be that you become *less confident* of your interpretation of what is attributable but still do not end up agreeing with your father's interpretation. That is one way in which learning about discord has some potential to moderate discord, but I have something stronger in mind in what follows.

be responding to. That means that it is shaped by what you interpret them as *thinking* or *perceiving* is attributable to the person they are responding to. But if you fail to recognize that you are in discord with them, then what you interpret them as thinking or perceiving is attributable to someone will be the same as what *you* think or perceive is attributable to them. Consequently, when you fail to recognize that you are in discord, charity can push you toward an attributability interpretation that is different from what charity would push them toward. So some discord is bound to persist precisely because it is unrecognized.

My suspicion is that this dynamic is often at least to some degree implicated when we have persistent disagreements about small matters concerning how to live together, such as where to keep the toothpaste, how to do the laundry, or how often to take out the trash. When people live together, they often have competing preferences or values about how such small things are handled. Ideally, much of the time, these preferences are weak enough or unstable enough that both parties can meet in the middle and find new ways of doing things that work for each of them. But this is not, I conjecture, how it always goes. We often experience one another's expectations about how such things are to be handled as *requests* that they are making of us, which we can only judge as reasonable to the extent that they are justified by benefits that exceed the force of our competing requests to do things in a different way.

When someone requests that you do things their way even though you have issued a competing request that they do things your way—or at least that you be allowed to—this tells us something about their priorities. It tells us that they think the values to be realized by their way of doing things are more important than honoring your request. At least, it tells us this if we assume that they attribute this request to you. For honoring a request is a participant response, and we do not honor requests that are not actually attributable to someone (such as an unwilling addict's request that you return her needles). But other things being equal, it is uncharitable to interpret one's life partner as caring more about (say) whether we keep extra folded trash liners at the bottom of the trash can to replace the old one when we take out the trash, than about honoring your request not to. So charity can lead you toward failing to attribute that request to them. Maybe you see it as pathological or compulsive. It paints a more positive picture of them to see them as having a hangup over the trash liner thing but caring a lot about honoring your requests, than as caring more about this very specific thing than about what you want.

But the charitability of this interpretation turns essentially on assuming that they attribute your request to you. And unfortunately, the situation is symmetric. They have also requested you to do something—and you are persisting in your request to do otherwise. They attribute their own request to

themselves, and so an interpretation of you has to decide whether you care more about whatever value (laziness, perhaps) is served by not having to follow their request than about honoring their request. From their perspective, *that* does not look charitable. And so from their point of view, it can look more charitable to interpret *your* request as not attributable to you.

Importantly, the charity of *each* of these interpretations turns on failing to recognize that the two of you are in discord. Once you recognize that the other person does not interpret what is attributable to you in the same way that you do, charity does not require failing to attribute their request to them in order to avoid taking them to care more about some minor household preference than about honoring your requests.

4.3. *Discord Clumping*

As the foregoing example illustrates, sometimes two or more cases of discord can be coordinated with one another in an interesting way. You and your partner are each making a request of the other, and each of you interprets your own request as attributable to you but is in doubt about whether the other's request is really attributable to them. The way that each attributability interpretation is sustained by charity offers an explanation of why the alignment of these two separate mismatches between what you interpret as attributable and what they do is not a coincidence. They come together because each helps to sustain the other.

The fact that attributability interpretation is value laden makes it likely that discord is *often* clumped like this. Because our behaviors are connected to our values in complex ways, it is often the case that when charity supports not attributing some particular behavior to someone, it also supports not attributing other closely related behaviors. For example, your friend has had one too many drinks at the end of the night and asks for their keys back. You decide that this is just the alcohol speaking and hang onto their keys. But then they get upset. "Don't you remember that paper about discord that we read last week? You're making an error in not attributing this decision to me—it's up to me to decide whether I'm sober enough to drive." Now they are trying to talk their way out of discord and using me to try to do it. But the very same reasons that support interpreting their request for the keys as not really attributable to them also support interpreting their insistence that it really is attributable to them as likewise not attributable to them.

5. CONCLUSION

In this article I have sought to introduce and explain the concept of discord. Discord builds on the concept of attributive responsibility, but because, as I

have argued, attributive responsibility marks a distinction that all of us mark as an integral of all of our interpersonal relationships, it is not just a philosopher's concept—it describes an ordinary phenomenon that can and does arise in ordinary interpersonal relationships. Indeed, I have argued that discord is unavoidable—we are bound to make some mistakes about what is attributable to someone, and, in particular, we are bound to sometimes underproject attributability. This leads us to fail to listen to, be grateful to, or be proud of one another even in cases in which doing so is licensed. I have also shown that the failure to recognize discord can have deleterious consequences of its own. It can help discord to persist longer than it would otherwise, and it can create illusions of ill will as we triangulate on others' motives and priorities from mistaken interpretations of what they are responding to.

All of these things are possible and have real consequences for interpersonal relationships, no matter what attributability really is and no matter how we actually think about it. But I have also argued that as a matter of fact, how we determine what is attributable to someone is filtered through our own values because we interpret through the lens of *charity*. And although it is not required for the further points that I have made in this article, we are *right* to do so. If this is right, then it can help us to understand many other important things about discord—including its prevalence. Value disagreements are everywhere. And if we count by its earmarks, discord is too.²⁶

University of Southern California
maschroe@usc.edu

26 The ideas in this article have benefited from audiences at UC San Diego, University of Pittsburgh, UC Santa Cruz, Stanford University, SUNY Buffalo, Syracuse University, University of Illinois, Australian National University, Colgate University, UCLA, University of Virginia, Harvard University, University of Edinburgh, St. Andrews University (twice), University College London, Southampton University, King's College London, University of Groningen, Ohio State University, University of Toronto, the Conceptual Foundations of Conflict Project at USC, the 2023 Mangoletsi-Potts Lectures at Leeds University, and the 2024 Moha Lectures at Seoul National University. The number of people to whom I am deeply indebted for discussions of this and related material is too great to exhaustively list here, but I want to call out for special notice Rima Basu, Steve Bero, Gwen Bradford, Susan Brison, Shishir Budha, Juliet Deng, Alex Dietz, Jan Dowell, Erik Encarnacion, Maegan Fairchild, Charlotte Figueroa, Carolina Flores, Mark Greenberg, Brian Haas, Eric Henney, Rianna Herzlinger, Fernando Rudy Hiller, Joe Horton, Nathan Howard, Jessica Isserow, Aaron James, Robin Jeshion, Zoë Johnson King, Renee Jorgensen, Rachel Keith, Amelia Khan, Matt King, Shieva Kleinschmidt, Barry Lam, Gerald Lang, Woo Ram Lee, Rebecca Lemon, Hilde Lindemann, Berislav Marusic, Nurit Matuk Blaustein, Mary Kate McGowan, Jessica Moss, Shyam Nair, Michael Nelson, Alexander Prescott-Crouch, Wyatt Radzin, Jon Quong, Lucy Randall, Grant Rozeboom, Marya Schechtman, David

REFERENCES

- Arpaly, Nomy, and Timothy Schroeder. *In Praise of Desire*. Oxford University Press, 2013.
- Beadle, Robert. "America's Top 10 Best-Selling Candy Bars of the Year." *Candy Retailer* (blog), November 3, 2024. <https://www.candyretailer.com/blog/top-10-best-selling-candy-bars/>.
- Bosco, Irene. *The Triangle of Innocence*. PhD diss., University of Southern California, forthcoming.
- Bratman, Michael. *Structures of Agency*. Oxford University Press, 2007.
- Dewey, John. *Outline of a Critical Theory of Ethics*. Hillary House, 1957.
- Dotson, Kristie. "Tracking Epistemic Violence: Tracking Practices of Silencing." *Hypatia* 26, no. 2 (2011): 236–57.
- Elga, Adam. "Reflection and Disagreement." *Nous* 41, no. 3 (2007): 478–502.
- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68, no. 1 (1971): 5–20.
- . *The Importance of What We Care About*. Cambridge University Press, 1988.
- Hesni, Samia. "Illocutionary Frustration." *Mind* 127, no. 508 (2018): 947–76.
- Holton, Richard. "Deciding to Trust, Coming to Believe." *Australasian Journal of Philosophy* 72, no. 1 (1994): 63–76.
- Knobe, Joshua, and Erica Roedder. "The Ordinary Concept of Valuing." *Philosophical Issues* 19, no. 1 (2009): 131–47.
- Korsgaard, Christine. "Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations." *Philosophical Perspectives* 6 (1992): 305–32.
- . *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press, 2009.
- Langton, Rae. "Duty and Desolation." *Philosophy* 67, no. 262 (1992): 481–505.
- . "Speech Acts and Unspeakable Acts." *Philosophy and Public Affairs* 22, no. 4 (1993): 293–330.
- Langton, Rae, and Jennifer Hornsby. "Free Speech and Illocution." *Legal Theory* 4, no. 1 (1998): 21–37.
- Lindemann, Hilde. *Damaged Identities, Narrative Repair*. Cornell University Press, 2001.
- May, Joshua, and Richard Holton. "What in the World Is Weakness of Will?"
-
- Shoemaker, Sam Shpall, Justin Snedegar, David Sosa, Aaron Suduiko, David Sussman, Lara Svirsky, Rob Tempio, Stephanie Van Fossen, Pekka Väyrynen, Vilma Venesmaa, Timothy Williamson, Levy Wang, Shu Wang, Shane Ward, Robbie Williams, Nadja Winning, Elise Woodard, and two anonymous referees and the editors of *JESP*.

- Philosophical Studies* 157, no. 3 (2012): 341–60.
- McGowan, Mary-Kate. *Just Words: On Speech and Hidden Harm*. Oxford University Press, 2019.
- Newman, George, Paul Bloom, and Joshua Knobe. “Value Judgments and the True Self.” *Personality and Social Psychology Bulletin* 40, no. 2 (2013): 203–16.
- Newman, George, Julian de Freitas, and Joshua Knobe. “Beliefs About the True Self Explain Asymmetries Based on Moral Judgment.” *Cognitive Science* 39, no. 1 (2014): 1–30.
- Phillips, Jonathan, Luke Misenheimer, and Joshua Knobe. “The Ordinary Concept of Happiness (and Others Like It).” *Emotion Review* 3, no. 3 (2011): 929–37.
- Phillips, Jonathan, Sven Nyholm, and Shen-yi Liao. “The Good in Happiness.” In *Oxford Studies in Experimental Philosophy*, vol. 1, edited by Tania Lombrozo, Joshua Knobe, and Shaun Nichols. Oxford University Press, 2014.
- Pizarro, David, Eric Uhlmann, and Peter Salovey. “Asymmetry in Judgments of Moral Blame and Praise: The Role of Perceived Metadesires.” *Psychological Science* 14, no. 3 (2003): 267–72.
- Resnick, Brian. “The Internet Peaked with ‘The Dress,’ and Then It Unraveled.” *Vox*, June 3, 2024. <https://www.vox.com/24117882/the-dress-blue-black-white-gold-internet-viral-media-perception>.
- Schechtman, Marya. *The Constitution of Selves*. Cornell University Press, 1996.
- Schroeder, Mark. “Attributive Silencing.” *Oxford Studies in Normative Ethics*, vol. 12, edited by Mark Timmons. Oxford University Press, 2022.
- . *Interpretive Objects: Meaning in Language, Life, and Law*. Unpublished manuscript.
- . “Narrative and Personal Identity.” *Aristotelian Society* 96, no. 1 (2022): 209–26.
- . “Persons as Things.” *Oxford Studies in Normative Ethics*, vol. 9, edited by Mark Timmons. Oxford University Press, 2019.
- . “Tipping Points: Abuse and Transformative Discovery.” *Free and Equal* 1, no. 1 (2025): 1–35.
- . *When Things Get Personal: The Hidden Philosophy of Conflict*. Unpublished manuscript.
- Shoemaker, David. *Responsibility from the Margins*. Oxford University Press, 2017.
- Strawson, Peter. “Freedom and Resentment.” *Proceedings of the British Academy* 48 (1962): 1–25.
- Watson, Gary. “Free Agency.” *Journal of Philosophy* 72, no. 4 (1975): 205–20.
- . “Two Faces of Responsibility.” *Philosophical Topics* 24, no. 2 (1996): 227–48.

WRONGING PERSONS THROUGH PHOTOGRAPHS ON PHOTOGRAPHIC FRONT INCURSIONS

Macalester Bell

There is something appalling about photographing people. It is certainly some form of violation.

—Henri Cartier-Bresson

Well, one of the reasons I photographed my family when I first started doing it was because I could push them around. I didn't feel uncomfortable about using them.

—Philip-Lorca diCorcia

MANY PHOTOGRAPHS take persons as their primary subjects, and these images raise a number of moral questions that have received little sustained philosophical attention.¹ People often feel wronged by the creation and dissemination of their photographic image. In fact, reports of such feelings are as old as photography itself, and the seemingly predatory nature of photography is baked into many English terms for photographic activity. For example, photographs are *shot*, *taken*, or *captured*. In some legal

1 Analytic philosophers who address the general ethical dimension of photographs of persons include Bell, "Respecting Photographic Subjects"; Davies, "Susan Sontag, Diane Arbus, and the Ethical Dimensions of Photography"; Hadley, "Street Photography Ethics"; and Danto, "The Naked Truth." Other theorists who have written about the ethical dimensions of photographs of people include Coleman, "Private Lives, Public Places"; Linfield, *The Cruel Radiance*; and Sontag, *On Photography* and *Regarding the Pain of Others*. While not focused on still photographs, Rini and Cohen offer an account of how so-called "deepfake" videos can harm ("Deepfakes, Deep Harms"); and Rini discusses how deepfake videos can undermine the role that audio and video recordings play in providing what she calls an *epistemic backstop* ("Deepfakes and the Epistemic Backstop"). In addition, privacy theorists have sometimes discussed photographic invasions of privacy as part of an overarching argument in favor of a particular conception of privacy. In fact, it was advances in photography (allowing for surreptitious picture taking) and concerns about nonconsensual dissemination of photographic portraits that inspired Warren and Brandeis' law review article "The Right to Privacy," which was the first publication in the United States to make the case for a legal right to privacy.

systems, taking or distributing a photograph without the express consent of the subject is a violation of their personality rights.²

Yet many dismiss these claims of purported wrongdoing or are quite conservative in their estimations of when photographs can wrong their subjects, e.g., by suggesting that photographs wrong primarily when they violate someone's privacy or portray offensive stereotypes. While these skeptics acknowledge that people can become emotionally upset by having their photograph taken or made unhappy by the resulting image produced, this unhappiness is often considered simply the result of an objectionable form of vanity or preoccupation with one's image. We do not, according to these critics, *owe* anyone happiness in this domain, and it is incorrect to suggest that we can wrong someone through the creation or display of a photograph, except in a narrow range of cases (e.g., photographs that violate a person's privacy).³ And the fact that many people are not at all bothered by being photographed and even welcome the opportunity further cements the impression that the vast majority of claims of wrongdoing in this domain are simply exaggerated expressions of bruised egos and can therefore be safely ignored.

I think these skeptical assumptions about the reality of photographic wrongdoing are mistaken. While it is possible, of course, to dislike a photograph of oneself due to vanity, and while we do not have a general moral obligation to ensure that every person's vanity is protected, photographs can do more than bruise people's egos. Photographs can wrong persons in ways that go beyond privacy violations, stereotyping, and other kinds of widely recognized wrongs, and we should take this possibility seriously in our photographic practices.

I start by briefly outlining two cases in which persons felt wronged by the creation, alteration, handling, or dissemination of a photograph of themselves. Specifically, I am interested in cases where the subject of the photograph has purportedly been wronged. These testimonies obviously do not settle the matter of whether or how a photograph can wrong, and we cannot accept at face value these testimonies as decisive or irrefutable evidence of wrongdoing. But given widespread skepticism regarding photographic wrongdoing and the general tendency to dismiss many claims of wrongdoing as mere vanity or overpreoccupation with one's image, beginning with these testimonies and attempting to make sense of them is methodologically important: when philosophizing about potential wrongs that are not widely recognized, we have good reason to begin with the testimony of those who take themselves or others to

2 For an overview of country-specific conceptions of personality rights, see the Wikipedia entry for "personality rights."

3 I am grateful to Daniel Star for this way of putting the critics' point.

have been wronged and to use that testimony as a starting point in thinking about the potential wrongs in question. A bottom-up methodology that begins with testimonies of purported wrongdoing and then attempts to make sense of apparent wrongdoing by interrogating motivating assumptions, connecting these claims to recognized forms of wrongdoing and so on, is important in a domain where the very existence of the wrong is questioned.

The methodology used in this article is bottom-up in an additional way: I aim to avoid imposing a specific normative framework on the terrain from the outset. Instead, I allow the testimonies to guide the moral theorizing.

My examples are cases involving photographs in particular, not pictures in general. While similar complaints can perhaps be raised against the use of drawings, etchings, and other static images on the one hand, or videos, films, or other moving images on the other, I hope to show that photographs can wrong persons in unique ways due to what I describe as the symbolic value of photographs and the specific role that photographs of persons play in our admittedly contingent practices. At the same time, I acknowledge that the line between photographs and other types of images is often blurry, so to speak, and insisting on a sharp distinction between photographs and other images is not necessary for my argument.

1. MOTIVATING EXAMPLES: FEELING WRONGED BY A PHOTOGRAPH

To begin, consider the following two real-life cases in which a person reported feeling wronged by the creation, handling, or display of a photograph.

Abigail Roberson: Abigail Roberson, born in 1883, was a teenager when she had her photographic portrait taken in a Rochester, New York, studio for circulation among her friends and family. Without Roberson's knowledge or consent, the resulting portrait was subsequently sold and eventually became central to an advertisement for Franklin Mills Flour. The image showed Roberson in profile with the words "Flour of the Family" emblazoned above. Twenty-five thousand advertising posters featuring Roberson's image were created and distributed to stores, warehouses, saloons, and other public places, including some in Roberson's hometown of Rochester. The effects on Roberson were profound, and she filed a lawsuit alleging that she had "been recognized by friends . . . and other people with the result that [she] has been greatly humiliated by the scoffs and jeers of persons who have recognized her face and picture on this advertisement, and her good name has been attacked, causing her great distress and suffering both in body and mind; that she

was made sick and suffered a severe nervous shock, was confined to her bed and compelled to employ a physician.”⁴

Erno Nussenzweig: For his photographic series “Heads” (2000–2001), the photographer Philip-Lorca diCorcia rigged a strobe light on scaffolding in Times Square and, with a very long lens, took photographs of pedestrians walking by. Subjects did not know they were being photographed, and diCorcia did not ask for their consent after the fact. The resulting images are imposing, large-scale prints that have been exhibited in galleries around the world. When Erno Nussenzweig learned that his portrait, *Head No. 13, 2000*, was being sold for between twenty thousand and thirty thousand dollars in a Manhattan art gallery, he sued diCorcia and the gallery for damages. In part, Nussenzweig objected to the portrait on religious grounds, claiming that as an Orthodox Hasidic Jew, the portrait violated his freedom to practice religion insofar as diCorcia used his image in what Nussenzweig saw as a violation of the prohibition against graven images. But Nussenzweig also objected that the photograph constituted a *failure of respect*. As his lawyer put it, Nussenzweig felt he had “lost control of his own image” and resented that his dignity had been compromised.⁵

These two examples highlight the persistence of claims regarding photographic wrongs, despite widespread skepticism about their very existence. Nonetheless, it is important to recognize that these cases are distinct along several dimensions. For example, Nussenzweig seemed to feel wronged by the *creation* of the photograph, whereas Roberson reportedly felt fine about the creation of the photograph but wronged by the handling and dissemination of the image.

Through their creation, handling, display, or dissemination, photographs can arguably wrong a person or group in (at least) the following ways:

- by lying about the subject;
- by being created through deception;
- by presenting the subject in an objectionable light;

4 *Roberson v. Rochester Folding Box Company*, 171 NY 538 (1902).

5 Gafter, “The Theater of the Street, the Subject of the Photograph.” This case is also discussed in Bell, “Respecting Photographic Subjects.” The legal case *Nussenzweig v. DiCorcia* was decided in favor of the defendant by the Supreme Court, New York County, on February 8, 2006. The court ruled that the statute of limitations had expired, and the photograph was a work of art protected by free expression under the First Amendment. This ruling was upheld by the New York Supreme Court Appellate Division in March 2007; in November of that year, the New York Court of Appeals upheld previous decisions.

- by presenting something private about the subject;
- by financially exploiting the subject;
- by inappropriately aestheticizing the subject;
- by inappropriately sexualizing the subject;
- by trivializing the morally serious;
- by symbolically attacking the subject.⁶

But what our examples of Roberson and Nussenzweig bring out is that there may be a distinct, further way that photographs can wrong. In the next section, I argue for the existence of a type of wrong that I call nonconsensual *photographic front incursion*. I take the language of *front* from Erving Goffman's dramaturgical account of social interactions. As I argue, we may wrong someone

- by creating, altering, handling, or disseminating a photograph of someone without the subject's informed consent and without also lying about them, deceiving them through the creation of the photograph, presenting them in an objectionable light, presenting something private about the person, financially exploiting the person, aestheticizing them, sexualizing them, trivializing something morally serious, or symbolically attacking them.

As the description above indicates, this type of wrong is best understood negatively: it is constituted by a kind of disrespect that is separate from (although it can contribute to and amplify) the nine other photographic wrongs outlined above. While there is arguably an element of disrespect in all ten types of photographic wrongdoing, disrespect is central to nonconsensual photographic front incursion.

We might think it is unnecessary to introduce the concept of photographic front incursions because the sense of being wronged inherent in the two motivating examples can be captured by other, more familiar normative categories such as privacy. In fact, in their lawsuits, Roberson and Nussenzweig both claimed to suffer violations of privacy from the creation and dissemination of their images, but there is something puzzling about their claims: the images in question do not seem to reveal anything intimate, personal, or private about them. I do not attempt to provide an overview of philosophical theories of the moral right to privacy here, but our motivating examples are not easily characterized as privacy violations. Consider, as an illustrative example, Carissa Veliz's gloss on privacy:

6 For further discussion of these ways that photographs can wrong a person, see Bell, "On the Variety of Photographic Wrongs."

Privacy is the quality of having one's personal information and "sensorial space" unaccessed. You have privacy with respect to a certain person to the extent that that person does not know anything personal about you, and to the extent they cannot see, hear, or touch you in contexts in which people do not commonly want to be the object of others' attention.⁷

If this is how we should understand privacy, then it is difficult to see how our motivating examples could be understood as privacy violations. Nothing especially personal or intimate is revealed through the photographs of Roberson and Nussenzweig: the first is a conventional studio portrait, and the second shows the subject walking down a street in midtown Manhattan.

Despite this, I think there is something to Roberson's and Nussenzweig's claims to have been wronged through the creation and dissemination of these images. I argue that they were both victims of nonconsensual photographic front incursion insofar as they were disrespected as self-presenting agents. This form of wrongdoing cannot be reduced to any of the nine types of wrongdoing listed above.

2. MAKING THE CASE FOR A FUNDAMENTAL PHOTOGRAPHIC WRONG: PHOTOGRAPHIC FRONT INCURSION

In his lawsuit against diCorcia, Nussenzweig objected that the photograph of him, *Head No. 13, 2000*, constituted a form of disrespect. As his lawyer put it, Nussenzweig felt he had "lost control of his own image," and he resented that his dignity had been compromised.⁸ I think there is something to Nussenzweig's complaint, and I hope here to build on what he expressed. As I see it, diCorcia disrespected Nussenzweig as a self-presenting agent and, in so doing, wronged him in a distinct way.

To preview, I argue that photographers, image handlers, and exhibitors have a special duty of respect toward their photographic subjects *qua* photographic subjects when their subjects are clearly identifiable persons. This respect requires that subjects give their informed consent before having a photograph of them taken, handled, altered, or disseminated. Photographers and other image handlers disrespect photographic subjects when they take,

7 Veliz, "Self-Presentation and Privacy Online," 35–36. Other influential accounts of the right to privacy include Allen, *Unpopular Privacy*; Warren and Brandeis, "The Right to Privacy"; Carnegie-Arbuthnott, "Privacy, Publicity, and the Right to be Forgotten"; Inness, *Privacy, Intimacy, and Isolation*; Rachels, "Why Privacy Is Important"; Scanlon, "Thomson on Privacy"; and Thomson, "The Right to Privacy."

8 Geffer, "The Theater of the Street, the Subject of the Photograph."

handle, or alter photographs of identifiable persons without subjects' consent. This duty of respect has two primary grounds: our status as self-presenters and the ways in which photographs function as natural symbols of their subjects and thereby offer a distinctive tool of extending agency.⁹ When this duty of respect is flouted, subjects are wronged insofar as they are disrespected as photographic subjects.

2.1. *Kantian Respect and Photographic Wrongs*

While I think the wrong of photographic front incursion is best conceptualized as a distinct form of disrespect, it is worth pausing to consider why a standard Kantian account of respect for persons has difficulty capturing and making sense of the way that Roberson and Nussenzweig were wronged.

Kant's second formulation of the Categorical Imperative enjoins us to "act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means."¹⁰ In *Groundwork of the Metaphysics of Morals*, Kant seems to identify "humanity" with rational nature. For example, he writes, "the mere dignity of humanity as rational nature, without any other end or advantage to be attained by it—hence respect for a mere idea—is yet to serve as an inflexible precept of the will."¹¹ In *The Metaphysics of Morals*, Kant writes that humanity is "the power to set an end . . . any end whatsoever."¹² According to Thomas E. Hill, we should understand the "humanity" that we ought to respect as the powers necessary for rationality and end setting. Specifically, humanity includes the capacity to act on maxims and to follow hypothetical imperatives.¹³ Humanity is thought to include a type of freedom that nonhuman animals lack, including "the ability to foresee future consequences, adopt long-range goals, resist immediate temptation, and even commit oneself to ends for which one has no sensuous desire."¹⁴ Humanity includes acceptance of categorical imperatives, as well as a basic ability to

9 In "Respecting Photographic Subjects," I argue that subjects whose photographs are taken without their consent are wronged insofar as they are forced to take up an alien (and likely distorted) perspective on themselves, which they never consented to. I go on to argue that subjects' status as self-presenters is compromised insofar as they lack agency over how they are depicted in the photograph. Here, I expand on this second point and offer a sustained account of how photographic front incursion undermines persons' agency, given photographs' status as natural symbols.

10 Kant, *Groundwork of the Metaphysics of Morals*, 4:429.

11 Kant, *Groundwork of the Metaphysics of Morals*, 4:439.

12 Kant, *The Metaphysics of Morals*, 4:392.

13 Hill, "Humanity as an End in Itself," 86.

14 Hill, "Humanity as an End in Itself," 86.

process information about the world.¹⁵ As Hill reminds us, for Kant, “physical abilities are not part of humanity in us, for humanity is identified with our noumenal personality as distinct from the phenomenal, or observable, person.”¹⁶

As is clear from this brief summary of the standard interpretation of *what* we are enjoined to respect when we respect persons, the object of Kantian respect is not the *personas* or *bodies* of persons; instead, we ought to respect the *dignity* of persons as rational agents. But if this is right, how can taking a photograph of a person ever be a failure of respect? After all, photographs are images of persons’ physical bodies, not their noumenal selves. In other words, if we accept the standard view of what we ought to respect when we respect persons, there does not seem to be conceptual space for making sense of why Roberson and Nussenzweig felt disrespected.¹⁷

But Kant’s picture is complicated by the fact that he explicitly acknowledges that humans are *embodied* rational agents.¹⁸ The fact that we are embodied rational agents means that we are constrained, morally speaking, in what we can do to our own (and others’) bodies:

If the body were related to life not as a condition but as an accident or circumstance so that we could at will divest our selves of it; if we could slip out of it and slip into another just as we leave one country for another, then the body would be subject to our free will, and we could rightly have the disposal of it. This, however, would not imply that we could similarly dispose of our life, but only of our circumstances, of the moveable goods, the furniture of life. In fact, however, our life is entirely conditioned by our body, so that we cannot conceive of a life not mediated by the body and we cannot make use of our freedom except through the body. It is, therefore, obvious that the body constitutes a part of ourselves.¹⁹

Given that our life is “entirely conditioned by our body,” to respect our rational agency, we must also, to some degree, respect our bodies (even though, as

15 Hill, “Humanity as an End in Itself,” 86.

16 Hill, “Humanity as an End in Itself,” 86.

17 Davies presses a version of this point in exploring whether the relation between a photographic subject and their photograph should be understood as analogous to that between real people who serve as models for fictional characters and the subsequent fictional works, which have become known as *reality fiction*. See Davies, “Susan Sontag, Diane Arbus, and the Ethical Dimensions of Photography.”

18 Several Kantians stress the moral importance of our embodied nature. See, for example, Korsgaard, *Fellow Creatures*; and Herman, “We Are Not Alone.”

19 Kant, *Lectures on Ethics*, 147–48.

Barbara Herman notes, Kant has “distain for the body”).²⁰ This is what grounds some of Kant’s more notorious claims, e.g., that selling our hair is not “entirely free from blame” and that sexual gratification outside of marriage involves a failure to respect oneself and one’s partner. For Kant, we risk throwing ourselves away (i.e., disrespecting ourselves) if we treat our bodies in degrading ways, even if we do so consensually.²¹ Many of the examples of bodily degradation that Kant gives involve selling one’s body or parts of one’s body or allowing someone to use one’s body for sexual satisfaction. These arguments seem to turn on some rather controversial assumptions about what sort of behavior is intrinsically degrading to human beings *qua* embodied rational agents, and it is not clear why, for example, selling one’s hair is verboten, but selling one’s labor through the marketplace is permissible.²²

Some contemporary Kantians insist that our experience of being embodied creatures ought to shape how we understand and discharge our duties toward others. For example, Herman argues that in order to fulfill our duties to support and cooperate with others, we must rely on our experience of different sorts of ends that we have as embodied agents, as well as on the ways our embodiment may impede us (e.g., through pain or other physical limitations). It is through our knowledge of what we are like as sensible creatures and our relationships with other embodied creatures that we are able to discharge our duties.²³

While the standard view should be refined in light of Kant’s remarks about embodiment, it remains unclear what a Kantian should think about photographic front incursions. Although Kant’s writings can perhaps provide resources for answering questions about, say, the ethics of selling human organs, we need to move beyond Kant’s texts to understand the purported photographic wrong under discussion here. While Kant does offer limited remarks about our status as embodied rational creatures, these remarks are too sparse and too controversial to support a clear diagnosis of what went wrong in the motivating examples of Roberson and Nussenzweig. What we need is further guidance on how to engage with one another as merely contingent rational agents.²⁴

There is a way of thinking about photographic front incursions as an important kind of disrespect. What we ought to respect when we respect persons includes the tools those persons need to exercise their agency; and these tools

20 Herman, “Could It Be Worth Thinking About Kant on Sex and Marriage?” 54.

21 Kant, *Lectures on Ethics*, 119.

22 For discussion of these issues, see Chadwick, “The Market for Bodily Parts”; and Tadd, “The Market for Bodily Parts.”

23 Herman, “We Are Not Alone,” 166–67.

24 I owe this way of putting the issue to an anonymous reviewer.

include a reasonable degree of control over their faces and photographic images of their faces. While my approach goes beyond Kant's limited discussion of what is involved in respecting persons as embodied agents in important ways, it is fundamentally Kantian in spirit. My approach stresses what is involved in respecting contingent rational agents who are embodied and who must live together in complex social environments.

2.2. *Disrespecting Persons as Self-Presenters: On Front Incursions*

As I see it, the photographs in our motivating examples disrespect the subjects as embodied self-presenting agents.²⁵ To respect someone as a self-presenter is to regard her as sovereign over her agential penumbra, or what Goffman calls her *front*. Photographers and image handlers disrespect their subjects insofar as they intrude on this agential space without consent (and often for self-serving reasons) through a photo-creation or handling process. Photographic subjects are *disrespected* insofar as their personas are treated as a mere means to others' ends. This kind of disrespect constitutes a wrong. In these cases, insufficient deference is given to the tools that persons need to exercise their agency.

Goffman's account of persons as self-presenters has been influential in normative ethical theory, particularly in theorizing the nature and moral importance of privacy and concealment. For Andrei Marmor, the right to privacy is grounded in persons' interests in "having a reasonable measure of control over the ways in which they can present themselves (and what is theirs) to others."²⁶ Yet Goffman's account of self-presentation has a wider application than has been heretofore recognized. As I argue, Goffman's account provides the resources for conceptualizing an undertheorized category of wrongdoing: what I call *front incursions*.

Using the metaphor of a theatrical performance, Goffman develops a dramaturgical model of society, arguing that we can understand social relations as different types of performances. Like actors on a stage, individuals perform various roles in their everyday social interactions. Through what Goffman calls *face-work*, we use our appearance, environment, and fellow actors to help us

25 Goffman introduces the notion of self-presentation in *The Presentation of the Self in Everyday Life*. Velleman develops Goffman's views in an interesting way in "The Genesis of Shame" and *How We Get Along*. Goffman's work also influences Nagel, "Concealment and Exposure"; Marmor, "What Is the Right to Privacy?"; Olberding, "Looking Philosophical"; and Berstler, "Bad Question!" Other philosophers appeal to Goffman's notion of *civil inattention* in order to address a range of philosophical issues. See, for example, Rini, "Contingency Inattention"; Gelfert, "Disattendability, Civil Inattention, and the Epistemology of Privacy"; Basu, "The Importance of Forgetting"; and Sharon and Koops, "The Ethics of Inattention."

26 Marmor, "What Is the Right to Privacy?" 4–5.

successfully put on a performance, while also showing a willingness to uphold the performances of others. Our goal in these interactions is to manage impressions. As Goffman puts it, “a person’s performance of face-work, extended by his tacit agreement to help others perform theirs, represents his willingness to abide by the ground rules of social interaction.”²⁷

Central to Goffman’s model is the distinction between the front of the house (i.e., the performance space) and the back of the house (i.e., the space in which actors are free from the gaze of the audience and can let down their guard). When a person meets others, she attempts to guide others’ impressions of her through a complex pattern of actions involving her setting and appearance.

Goffman’s dramaturgical model is most clearly applicable to social spaces in which persons have explicitly defined roles, but we also engage in impression management in public spaces. According to Goffman’s account of *civil inattention*, we show an apt level of indifference to strangers we encounter in public spaces. Rather than stare outright at others or ignore their presence completely, we show others that we acknowledge their existence without burdening them with excess attention.

Little philosophical attention has been paid directly to Goffman’s conception of the front.²⁸ The front is the “expressive equipment” that performers must utilize to conduct their performance.²⁹ Following Goffman, we can distinguish two parts of the front: the setting and the personal front. The setting includes the decor, furniture, and other props that actors must utilize to create their performances. The setting usually stays put, and for this reason, it marks the spatial boundary of the performance.³⁰ The personal front, on the other hand, is comprised of those elements of expressive equipment that travel with performers (e.g., insignias, clothing, looks, postures, facial expressions, bodily gestures, and so on).³¹ As should be clear from Goffman’s list, the face plays a central role in an individual’s personal front. Not only do our faces individuate us, but to a large degree, we exercise our self-presentations *through* our faces; and this is why photographs of identifiable faces raise special ethical concerns.

27 Goffman, *Interaction Ritual*, 31. For a philosophical discussion of the role of poise in face-work, see Berstler, “Bad Question!”

28 One exception is Olberding, “Looking Philosophical,” but Olberding focuses on “the existential dislocation that arises when what my stuff says to me is at odds with what it says to others” (693), not on how persons can be wronged by the incursions of others into their “stuff.”

29 Goffman, *Interaction Ritual*, 22.

30 Goffman, *Interaction Ritual*, 22.

31 Goffman, *Interaction Ritual*, 23–24.

When a person is unable to exercise agency over her personal front or loses control through “unmeant gestures” or slips, this often leads to feelings of helplessness and shame.

While philosophers influenced by Goffman have brought out the interest that persons have in self-presentation, an important distinction has been elided. Our interest in having a reasonable level of control over our self-presentations is multidimensional, and there are at least two different types of control (and two types of corresponding interest) that are relevant: we can distinguish between having an interest in what I call *front/back boundary control* and having an interest in *front control*. We exercise front/back boundary control over what, precisely, is concealed or hidden from view of the audience. Assuming that Marmor’s argument is broadly correct, we have an interest in having a reasonable level of control over what we hide and reveal about ourselves to different groups of people.³² Paradigmatic privacy violations seem to violate this interest that we have in keeping some aspects of ourselves hidden from at least some others. Front control, on the other hand, is not about control over the boundary between the front of house and the back of house, or what persons choose to reveal or keep hidden from the audience. Instead, front control is focused on the expressive equipment we need as self-presenters, and it includes control over our setting and personal front.

If we think that persons have an interest in having a reasonable level of control over what they hide and reveal about themselves due to their status as self-presenters, we should also conclude that persons have an interest in having a reasonable level of control over our expressive equipment. This interest is distinct from privacy interests. To be a self-presenting agent is to be someone who has some control over her self-presentations and appearance; for someone to discover that they had bright red lipstick on their teeth after delivering a lecture is embarrassing because it is evidence that they lacked this kind of agency, at least momentarily.³³ Critically, to discover that you have lipstick on your teeth is *not* to discover that your privacy has been violated. Instead, it is to discover a momentary loss of front control. In the cases that are of interest

32 Marmor, “What Is the Right to Privacy?” 7. For his part, Marmor elides this distinction between front/back boundary control and front control, preferring to talk in terms of control over self-presentations more generally. While he seems to be primarily focused on front/back boundary control, he pivots to a focus on front control in his brief discussion of street photography at the end of his essay. In not distinguishing between front/back boundary control and front control, Marmor’s framework ends up reducing front incursions to privacy violations. As a result, his account of privacy is vulnerable to objection. See Veliz, “Self-Presentation and Privacy Online” for some objections to Marmor’s account.

33 For Velleman, this discovery is shameful.

to us here, someone fails to respect the agent's authority over her own front. Our interest in front control is important in itself, and it is also a precondition for front/back boundary control; if agents lack front control, they are unable to exercise front/back boundary control.

Caring about ones' status as a self-presenter should not be conflated with being vain.³⁴ A vain person thinks that his appearance merits high esteem from others; threats to his vanity such as a bad hair day might leave the vain person without a secure basis for self-esteem. But threats to one's status as a self-presenter are fundamentally different. Having apt agency over one's self-presentations does not merit esteem but a basic kind of recognition respect. Significant threats to one's status as a self-presenter undermine one's status as a person.³⁵ It is telling that Roberson reported feeling *humiliated* by the dissemination of her photographic portrait; humiliation is precisely the emotion we expect when someone's status as a self-presenter is threatened. As Goffman puts it, when there is a contradiction between who a person has avowed themselves to be and their performance, they are vulnerable to "immediate humiliation and sometimes permanent loss of reputation."³⁶

Of course, we should not confuse sovereignty with complete control; no one has complete control over how they appear to others. (Many items included in Goffman's personal front are, in fact, impossible to control.) Nor is this a goal worth striving for.³⁷ As Marmor notes, it is *good* that we lack complete control over our self-presentations: "Nobody should have too much control over the way they present themselves to others, as that would make manipulation, dishonesty, and generally, lack of authenticity, all too easy.... [But people do need] some reasonable amount of control over the ways in which they present different aspects of themselves to others."³⁸ It is also good that we lack complete control over our fronts. We often inadvertently reveal things about our character through our slips. A perfectly curated self-presentation raises questions about the authenticity of the person in question. But we need to be able to exercise at least *some* agency over our front to be intelligible as a person.

34 Bell, "Respecting Photographic Subjects," 296.

35 In "The Genesis of Shame," Velleman writes, "You thus have a fundamental interest in being recognized as a self-presenting creature, an interest that is more fundamental, in fact, than your interest in presenting any particular public image. Not to be seen as honest or intelligent or attractive would be socially disadvantageous, but not to be seen as a self-presenting creature would be socially disqualifying: it would place you beyond the reach of social intercourse altogether" (37).

36 Goffman, *The Presentation of Self in Everyday Life*, 59.

37 Marmor, "What Is the Right to Privacy?"

38 Marmor, "What Is the Right to Privacy?" 8.

Part of what is involved in respecting another as a self-presenting agent is respecting the authority she has over the tools she needs to exercise this agency. When someone creates or takes my image—especially an image of my face—and uses it for their purposes without my consent (or if the power differentials between us make it impossible to give my informed and uncoerced consent), they are not showing appropriate deference to the tools I need to exercise this agency, and because of that, they are disrespecting me as a self-presenter.

Admittedly, some people do not seem to care very much about having a reasonable level of control over their expressive equipment. Given this, we might wonder if there is actually an interest here worth taking seriously. In response, it should be acknowledged that many people do not seem especially concerned about maintaining front/back boundary control either. But we should not take this lack of care as a reason to conclude that there is no genuine interest in privacy. Instead, we ought to see this lack of concern as a regrettable failure to recognize a genuine human interest. Similarly, I do not think the fact that some people do not seem to care much about front control shows that there is no interest here worth protecting.

To sum up, I argue that persons have an interest in being able to exercise reasonable control over their self-presentations, but this interest should not be understood exclusively in terms of controlling the front/back boundary. We also have an interest in being able to exercise a reasonable level of control over our fronts. Respect for a person *qua* self-presenter should therefore include the recognition that persons have authority over their fronts; failing to recognize this authority evinces disrespect for persons as self-presenters. Photographic front incursion involves a failure to recognize that persons *qua* self-presenters have authority over the tools they need to exercise their agency. We should not use another person's front for our own aims and projects without her consent. Taking, manipulating, or displaying a photograph of an identifiable person without that person's consent constitutes an incursion on their front and is, for this reason, morally wrong.

Front incursion is a general and undertheorized type of wrongdoing that may not have any special connection to photographs. I have offered only a sketch of the wrong of front incursion here; I hope to offer a more sustained articulation on another occasion.³⁹ But presumably, one can show this type of

39 Carnegie-Arbuthnott argues that violations of privacy, distortion, and defamation can each be understood as picking out separate interests collected under the general category of the historic right of personality, which aims to protect our interest in our reputations ("Privacy, Publicity, and the Right to Be Forgotten"). In the future, I aim to explore whether front incursions should be understood using this taxonomy—i.e., as a distinct type of wrongdoing, related to but distinct from violations of privacy, distortion, and defamation.

disrespect for persons in a number of ways—e.g., by forcing prisoners of war to wear the uniforms and insignia of opposing forces, or by the unauthorized use of a person's office and letterhead.⁴⁰ I have yet to make the case for the existence of *photographic* front incursions. If we accept that front incursions are a type of wrongdoing, why should we think that the taking, handling, or dissemination of a photograph of a person without their consent constitutes a photographic front incursion? I turn now to this question. I argue that we can and regularly do extend our agency through photographs, since (some) photographs function as natural symbols of persons; because of this, photographs of a person are an important aspect of that person's front.

2.3. *Photographs as Part of a Person's Agential Penumbra*

We can distinguish between *conventional* and *natural* symbols. In considering the differing grounds for offense in flag desecration and mistreatment of dead bodies, Joel Feinberg appeals to this distinction as follows:

A flag is an arbitrary or conventional symbol of an abstraction, which bears no striking similarity to what it symbolizes. Rather it comes to represent a country only by virtue of a conventional understanding. . . . A dead body, on the other hand, is a natural symbol of a living person and needless to say has a striking similarity to the real thing.⁴¹

While photographic images are not discussed in this context by Feinberg, I contend that some photographs are *natural symbols* of persons in Feinberg's sense, given their close resemblance to the photographic subject.⁴² That is, photographs can bear a "striking similarity" to their subjects, and a photograph in which a specific person is centrally featured and recognizable is a natural symbol of that particular person. Of course, just as with corpses, not all photographs closely resemble their subjects; but in cases where this resemblance holds, the photograph of a specific person is a natural symbol of the person. Moreover, these photographs are clearer natural symbols of persons than paint-

40 For the former, see Rainsford, "Ukraine War." Regarding the second example, recall the incident on January 6, 2021, when Richard "Bigo" Barnett broke into Nancy Pelosi's office, put his feet up on her desk, used some of her stationary to write her a note (addressing her with a sexist slur), and then bragged to people outside the Capitol that he "took Nancy Pelosi's office."

41 Feinberg, *Offense to Others*, 55.

42 Why and how some photographs closely resemble their subjects is beyond the scope of this essay. For discussions of the epistemic and "contact" power of photographs, see Walton, "Transparent Pictures"; Hopkins, *Picture, Image, and Experience*; and Cohen and Meskin, "On the Epistemic Value of Photographs."

ings, insofar as they more closely resemble their subjects than most paintings.⁴³ More specifically, it seems to me that photographs that resemble their subjects can be what we might call both *generic* and *specific* natural symbols. DiCorcia's photograph of Nussenzweig, for example, is a natural, specific symbol of Nussenzweig himself, and depending on the context, it could also be a natural, generic symbol of Orthodox Jewish men, or of persons more generally.

Interestingly, some professional photographers spontaneously describe the persons they photograph as symbols. Consider the following two examples:

I always prefer to work in the studio. It isolates people from their environment. They become in a sense—symbolic of themselves. I often feel that people come to me to be photographed as they would go to a doctor or fortune teller—to find out how they are.⁴⁴

I love the people I photograph. I mean, they're my friends. I've never met most of them or I don't know them at all, yet through my images I live with them. At the same time, they are symbols. The people in my pictures aren't Mr. Jones or Mr. Smith or whatever; they're someone that crossed my path or I've crossed their path, and through the medium of photography I've been able to make a good picture of that encounter. They have a life of their own, but they are also are symbols [sic]. I would say that I respect the viewer, but I don't want to tell him everything. Hopefully, there's an element of mystery involved. I like him to look at a picture and say "Well, that that reminds me of someone [sic]" and make up a little story in his head, make him smile, brighten up his day. I think this is what I'm trying to achieve with my photographs.⁴⁵

Because natural symbols closely resemble what they symbolize, we tend to be especially emotionally troubled when these symbols are affronted. As Feinberg notes, people tend to react more strongly to, say, corpse mutilation than to flag desecration, and the reason for this turns on the distinction between natural and conventional symbols: "when one mutilates a corpse, one is doing something that looks very much like mutilating a real person, and the spontaneous horror of the real crime spills over on the symbolic one."⁴⁶ The language of

43 There are, of course, some drawings that are nearly indistinguishable from photographs (e.g., Chuck Close's photorealist portraits). Insofar as these images are indistinguishable from photographs, my arguments apply to them as well. It is also worth noting that Close's portraits are based on photographic images.

44 Richard Avedon, quoted in Sontag, *On Photography*, 187–88.

45 American Suburb X, "An Interview with Bruce Gilden."

46 Feinberg, *Offense to Others*, 55.

“spilling over” suggests that Feinberg may not take affronts to symbols seriously or see them as genuine affronts. In fact, Feinberg is somewhat ambivalent about the value of respecting symbols. He does explicitly say that it is important to respect certain symbols, but he worries that we can get so caught up in valuing a symbol of something that we risk neglecting or otherwise undermining the value symbolized.⁴⁷ For our purposes, what is helpful about Feinberg’s discussion of symbolic affronts is that it allows us to make sense of the possibility that we could wrong persons through the creation, dissemination, or handling of photographs. The strong resemblance between some photographs and some photographic subjects gives photographs a distinctive symbolic power.

Photographic natural symbols of a person are part of the agential penumbra of the person. For one important way that persons extend and exercise their agency is through the natural symbols of photographs. We do things like *endorse* a product or political candidate by attaching our photographic image to an ad campaign. We also *claim* a particular social media profile as our own using photographic images. We *identify* ourselves as the author of a book through an author photograph. We *socialize* on some online platforms by navigating a photographic avatar through virtual space. This agential extension through photographs may explain why predatory leaders such as Mao Zedong and Warren Jeffs often display photographs of themselves throughout the physical spaces of their organizations; these photographs serve to remind their followers of their presence and authority. Given these practices of extending agency through photographic image, we are vulnerable to wrongs such as misrepresentation through deepfakes or spoofing (e.g., using images of other persons to create fake accounts on social media).

47 “Granted that it is important that we respect certain symbols, it is even more important that we do not respect them too much. Otherwise we shall respect them at the expense of the very values they symbolize and fall into the moral traps of sentimentality or squeamishness” (Feinberg, *Offense to Others*, 72). Further, “sentimental actions very often are excessive responses to mere symbols at great cost to genuine interests, one’s own or others’. In the more egregious cases, the cherished symbol is an emblem of the very class of interests that are harmed, so that there is a kind of hypocritical inconsistency in the sentimental behavior. William James’s famous example of the Russian lady who weeps over the fictitious characters in a play while her coach man is freezing to death on his seat outside the theater is an instance of sentimentality of this kind. The error consists of attaching a value to a symbol and then absorbing oneself in the sentiments evoked by the symbol at the expense of real interests, including the very interests the symbol represents. The process is not consciously fraudulent, for the devotion to sentiment may be sincere enough. Nor does it consist simply in a conflict between avowal and practice. Rather the faulty practice is partly *caused* by the nature of one’s commitment to the ideal. Sentimental absorption in symbols distracts one from the interests that are symbolized” (75).

As I argue, part of what is involved in respecting a person as an autonomous agent is respecting the tools she needs to exercise her agency; to respect a person as an agent is to respect the authority she ought to have over the tools she needs to exercise her agency. When someone intentionally creates, disseminates, or alters a natural symbol of another person without that person's consent, they disrespect that person as a self-presenter. While taking this attitude toward a person is at the heart of the wrong, the resulting image constitutes a material souvenir of this wrongful attitude—a souvenir that can be reproduced and circulated, which shows an even deeper level of disrespect.

Some might object that diCorcia did not wrong Nussenzweig in taking *Head No. 13, 2000* because not everything associated with an individual's persona is part of the person's front: only those things that are expressive of persons' attitudes or role are properly included in their front. And in *Head No. 13, 2000*, Nussenzweig is not expressing any attitude at all, nor does his comportment suggest a particular role identity. Therefore, it may be objected, Nussenzweig's self-presentation has not been affected by diCorcia's photographing him.⁴⁸ While it is true that Nussenzweig does not seem to have a specific attitude or expression on his face, nor is he expressing a clear role identity in *Head No. 13, 2000*, I do not think that settles the matter of whether he has been wronged. As I emphasize, we *do* things with our personas: we use them to endorse, protest, criticize, stand behind, and so on. To be able to perform these activities, we need exclusive authority over our front and over photographic images of our front. It might well be a further wrong to use an image of someone expressing a particular attitude or displaying a specific role identity, but just taking and using the image itself is wrong.

As my response to this objection indicates, I do not claim (or deny) that Nussenzweig and Roberson were in fact *harmed* by the consequences of having their photographs printed and displayed. Instead, what I want to say is that to see and to treat an important part of another person's agency as a mere resource for one's own projects is to wrongfully disrespect the other, given our status as self-presenters and given the way that photographs, as natural symbols of persons, are used to extend our agency. Even if Nussenzweig and Roberson had never become aware of the existence of their photographic portraits, they still were disrespected through the attitudes and acts of the photographers involved.⁴⁹ In both cases, those who took and disseminated the photographs failed to acknowledge the authority that Nussenzweig and Roberson should

48 I am grateful to Alan Patten for raising this objection.

49 Bell, "Respecting Photographic Subjects," 297.

have had over their fronts, and it is this attitude of disrespect that is constitutive of the wrong of photographic front incursion.

This is not to say that the actual consequences of photographic interventions never matter. Intuitively, it seems much worse to take a photograph of someone without their consent and publish it on a highly visible social media platform than to take the same photograph and keep the image private or share it only with a few friends. And Roberson was explicit that her distress arose because of the ubiquity of the advertising posters bearing her image. But my attitude-based account of the wrong of photographic front incursion can make sense of this. The wrong inheres in the disrespect shown, but when photographic images are widely released, the attitude expressed is different and worse, morally speaking. Not only is the photographic subject disrespected as a self-presenter, but given that photographs are so easily circulatable and given that being wronged in this way is often connected to feeling humiliated, widely disseminating the image worsens the wrong insofar as it evinces an attitude of *complete* lack of concern for the vulnerabilities of being a self-presenter.

If I am right that wronging in this domain primarily attaches to a disrespectful attitude, does it follow that the creation, handling, or display of a photograph can be justified in the absence of the subject's consent so long as one is careful not to have a predatory attitude toward the subject and her agency? I do not think so. The way that we avoid having a predatory attitude is by garnering consent for the creation or use of the image; we cannot avoid the charge of predation by having good ends or by prefacing our use of another's image simply by saying "I mean no harm" or "I do not wish to offend."

As should be clear, the way to avoid photographic front incursion is to obtain the permission of photographic subjects; ideally, the resulting image is a product of collaboration between subject and image creator.⁵⁰ However there are many barriers to gaining informed consent in this domain. A significant power differential between photographer and subject may make it impossible for the two to collaborate. Richard Avedon's 1969 photograph *Andy Warhol and Members of the Factory, New York City* may offer an instance of this phenomenon: one of the subjects, Candy Darling, was, according to reports, in awe of Avedon and the circles he traveled in, and because of Darling's entrancement with Avedon, Arthur Danto argues that she allowed him to take a brutal image of her, which she would not otherwise have endorsed.⁵¹ Even if we resist Danto's interpretation of this photograph, there is no doubt that Avedon seemed to take pride in his power over others. Being starstruck may make it impos-

50 Bell, "Respecting Photographic Subjects."

51 Danto, "The Naked Truth."

sible to give informed consent. And in a very different way, mental illness or cognitive limitations may also constitute a barrier to giving informed consent. Adam Broomberg and Oliver Chanarin's *Ghetto* is a series of photographs that depict people living in "gated communities," broadly conceived.⁵² Some of the photographs feature patients at a psychiatric hospital. While Broomberg and Chanarin are quite self-conscious about the power differential between photographer and photographic subject, and they play with this power imbalance in this series by allowing the patients to trip the shutter themselves with a long release cable, it is clear that many of the subjects do not fully understand the project or what role their images play in it.⁵³

As I note at the end of the previous section, I do not mean to suggest that *photographic* front incursions exhaust the category of front incursions. There are, I concede, other ways of disrespectfully intruding on a person's front.⁵⁴ Not only could other visual mediums be used in front incursions (e.g., line drawings) but presumably so could nonvisual mediums like *romans à clef* or well-publicized gossip. As Samuel D. Warren and Louis D. Brandeis note in their groundbreaking work on privacy, "If you may not reproduce a woman's face photographically without her consent, how much less should be tolerated the reproduction of her face, her form, and her actions, by graphic descriptions colored to suit a gross and depraved imagination."⁵⁵ Of course, some work needs to go into showing precisely how these other activities constitute front incursions, but nevertheless, front incursions can take a variety of forms. Given this concession, why do I specifically focus on photographic front incursions here? First, photographic front incursions are more common than pen-and-ink front incursions or front incursions through *roman à clef*, and their relative ubiquity merits special attention. Second, in recent years, the public has become attuned to purported photographic wrongs such as "revenge porn," "digital blackface," and "sharenting"; and an account of photographic front incursions aids our understanding of the wrong-making features of these activities. Third, as I argue in this section, photographs serve a special function in how persons extend their agency, especially in online environments. Because of that, photographs are especially important features of our agential penumbra. This feature of photographs is of course contingent: there is a possible world

52 Broomberg and Chanarin, *Ghetto*.

53 Of course, if subjects are seriously cognitively disabled, they may not be self-presenters at all. In this case, we need another account of what is morally problematic about these photographs.

54 I am grateful to an anonymous referee for pushing me to say more about this point and suggesting the example of *roman à clef* as front incursion.

55 Warren and Brandeis, "The Right to Privacy," 214.

in which we use etchings or pen and ink drawings as our avatars rather than photographs. But in the actual world, there is a tighter connection between a person and a photographic likeness than there is between a person and an ordinary line drawing.

To sum up, we have a fundamental interest in having a reasonable level of control over our expressive equipment, including photographic likenesses. To date, philosophers have not acknowledged this distinct interest and have tended to conflate this interest with our interest in privacy. Front incursions are wrongful violations of the authority that persons should have over their fronts. This kind of incursion manifests disrespect for persons *qua* self-presenters. In some cases, those who engage in front incursions evince a predatory attitude toward the fronts of others. Some might say that diCorcia manifested this type of predatory disrespect through his photographic practices.⁵⁶ However, in many cases, the disrespect shown through front incursions is more inconsiderate than predatory.

When it comes to privacy, most theorists think we have a *right* to privacy, not simply an interest in it. Do we similarly enjoy something like a right to front authority, which is wronged in cases of front incursions? I believe we do, and this right to front authority is what should undergird personality rights and can shed light on a number of legal controversies surrounding the use of photographic likenesses. However, I save a full articulation and defense of this right for another occasion. What I show here is that we do, in fact, have an *interest* in having a reasonable level of control over our expressive equipment, and if we have a right to something like front authority, this right is grounded in our interest, *qua* self-presenters, to have a reasonable level of control over our fronts.

3. OBJECTIONS AND REPLIES

In completing my defense of the wrong of photographic front incursion, I consider and respond to two objections here. First, some may resist my arguments by pointing to cases in which taking, handling, or disseminating a photograph of someone without their consent seems obligatory or at least permissible. Consider, for example, the wide and nonconsensual release of photographs of Justin Trudeau in blackface in the years before his prime ministership.⁵⁷ Let

56 Some might say that the TikTok creator Harrison Pawluck also displays this type of predatory disrespect toward persons' fronts. See, e.g., Touma, "Melbourne Woman 'Dehumanized' by Viral TikTok Filmed Without Her Consent." I am grateful to an anonymous referee for suggesting this type of example.

57 I am grateful to Carlos Santana for suggesting this example of Trudeau and raising this worry about my position.

us assume for the sake of the argument that appearing in blackface is morally wrong, and it is in the public interest for reporters to report upon the serious wrongdoing of public figures. While we could learn that Trudeau routinely dressed in blackface from the testimony of witnesses, seeing photographs adds something that such testimony lacks. Arguably, seeing photographs allows us to better appreciate the wrong of appearing in blackface. For this reason, we might think distributing the photograph of Trudeau without his consent was permissible or even obligatory. This case and others like it put pressure on my claim that nonconsensual photographic front incursion is a genuine wrong.

I concede that taking, handling, or displaying a photograph without a subject's consent might be justified if doing so is a necessary means to protecting the subject's interests or prevent another wrong to her or others.⁵⁸ That is, I am conceiving photographic front incursion as a kind of *pro tanto* wrong; invading the front of another is wrong absent special justification. In other words, this type of incursion may be necessary to protect the subject's interests in an emergency, or this type of incursion may be necessary to prevent a wrong to the subject or to another person. According to this line of thought, the police may be justified in disseminating a photograph of a missing person without first asking for her consent, or, perhaps more controversially, an antipornography activist might be justified in showing pornographic images as part of a campaign aimed at ending abuse in the porn industry. When it comes to preventing wrongs to others, we might think we are justified in using someone's photograph without their consent in putting out an all-points bulletin, for example, in cases where we have good evidence that someone poses a future imminent risk to the public.

Turning back to the example of the photographs of Trudeau in blackface, clearly their dissemination does not protect Trudeau's interests. Perhaps in a slightly different version of the example, it could be argued that the release of the images was a necessary means to prevent future wrongdoing. Yet in the actual case, it seems unlikely that the wrongful behavior in question was going to continue in the future; odds are that Trudeau the prime minister would

58 As part of their argument defending the wrongfulness of observing humiliating acts, Frowe and Perry argue that observation might be justified in cases where the observation benefits the victim. These include cases where observation is a necessary means of ending ongoing wrongdoing ("Wrongful Observation," 134). But they also suggest that observation might be justified in a wide variety of cases: "Observation may, for example, enable one to corroborate the victim's evidence in a subsequent prosecution, to show solidarity with the victim, or ensure that she will not be forgotten" (135). Although they go on to say that, in these cases, observation wrongs the victim if she has made it clear that she does not want to be observed. In addition, they assert that even if observation wrongs the victim, it does not follow that it is unjustified; sometimes we may permissibly infringe someone's right to non-observation in order to bring about greater benefits to others (135).

not continue dressing up in blackface. If this is right, then I would like to push back against the claim that the release of these images was permissible or even obligatory. While I can feel the pull of the intuition that members of the public are owed access to images of past wrongdoing of public figures, I think we ought to resist this idea. While the public may have a right to know about past wrongdoing that is relevant to a public figure's role responsibilities, we do not have a general right to *photographic* access to these wrongdoings. In this case, I believe that disseminating the photographs without Trudeau's consent was a wrongful photographic front incursion.

Second, many will balk at the claim that respecting photographic subjects requires that we always garner their informed consent because accepting this claim requires us to significantly alter our practices in ways that threaten other important values; for these reasons, the position defended here might seem overly demanding.⁵⁹ Would a photographic subject have a basis to complain if, for example, she attended a protest and a photograph of her participating in the protest is published the next day in the local paper without the photographer getting her consent?⁶⁰ In this case, the purported rights of the photographic subject seem to be directly at odds with the public's right to know about current events.

More generally, if the arguments of this article are sound, we also have reason to criticize much street photography.⁶¹ While some street photographers do seek consent from their subjects before taking photographs, many do not. And even those photographers who ask for consent often do not attempt to garner fully informed consent that would allow the subjects to see exactly how the photographer intends to use their image. Questions can also be raised about the ethics of seemingly innocuous activities such as collecting photographic portraits.⁶² The very act of curating a photographic collection involves quickly sorting through large numbers of photographic images, keeping those that fit

59 Some insist that we have a right to photograph that is grounded in our right to be informed. See Martinez and Renteln, "The Human Right to Photograph." Yet if the arguments of this article are sound, they seem to conflict with this purported right to photograph.

60 As an anonymous reviewer points out, the example involving photographing protesters is complicated by the fact that protesters are presumably self-consciously engaged in active face-work in a way that a pedestrian walking down a city street is not. Because of this, taking a photograph of protesters without their express consent seems like less of an incursion. Still, there is some reason to worry even here since the resulting images could be misrepresentative of what the subjects wish to convey. Additionally, in some contexts, being photographed at a protest could open the photographic subjects up to retaliation.

61 See Hadley, "Street Photography Ethics."

62 I owe this example to Julia Driver. Similar concerns can be raised regarding posting photos on online photographic collection sites such as <https://internetkhole.com> and <https://awkwardfamilyphotos.com>.

the collection and discarding those that do not. Does it follow that a collector who sorts through photographs in this way wrongs the subjects? In these and other cases, avoiding the apparent wrong of disrespecting a photographic subject seems to come at the expense of other values such as the public's right to be informed, the aesthetic value of candid street photography, and the value of seemingly innocuous hobbies like collecting photographic images.

As I have characterized photographic front incursion, at the heart of it is a certain *attitude*: persons are disrespected as self-presenting agents, and their personas are treated as mere means to others' ends. This attitude can be expressed in all these examples. So according to my argument, it is possible to wrong persons through these activities. However, there are many other values that are in play in these cases. There are aesthetic values, historical values, personal project values, and a person may, in a given instance, have overriding reason to create, manipulate, or disseminate photographic images of persons without their consent, despite the wrong done to photographic subjects.

While I do not have a recipe for how to balance the conflicting values at work in these cases, providing guidelines for resolving these conflicts is not really my aim here. Instead, my goal is to make the case for the existence of a type of wrong—nonconsensual photographic front incursion—that has not yet received philosophical attention. The account can help to make sense of certain attitudes that photojournalists, street photographers, and collectors may experience: the internet is full of advice aimed at helping street photographers and other photo handlers to get over their feelings of anxiety and trepidation about taking photos of people on the street, but if I am right, there is good reason for feeling uneasy when engaged in these activities. These feelings of discomfort may be accurately tracking the moral terrain; if my arguments are sound, these seemingly innocuous activities involve the commission of an unrecognized *pro tanto* wrong of photographic front incursion. While these activities might turn out to be all-in justified, depending on how we weigh competing values, my arguments suggest that there is an important moral remainder that merits our consideration.

It may be objected that if my arguments go through, I have set the bar for wronging someone as a photographic subject far too low. Many of us live in environments where we are constantly being photographed. Security footage is often taken of us without our knowledge or consent. To conclude that we are constantly being wronged by this activity seems to stretch the notion of wrongdoing too far. In response, I stress that the wrong of failing to respect someone as a photographic subject involves more than simply the creation of a photograph. Rather, failing to respect someone as a photographic subject involves one person (the photographer or image handler) taking or using an

image of another person to express the former's ideas without any collaboration with the subject. In the case of security cameras, there is no agent who is using another person's appearance in this way.

4. CONCLUSION

In this article, I have argued for the existence of a distinctive type of a general photographic wrong: nonconsensual photographic front incursion. If we acknowledge the existence of photographic front incursions, many further questions need to be answered: Are photographs of groups of people *less* likely to involve front incursions, or do these photographs pose the same problems outlined in this essay? What should we think about photographs of people whose faces are not clearly visible? Do we have any reasons to object to photographs of persons who are *not* self-presenters and who cannot consent to being photographic subjects, such as the severely cognitively disabled?

To date, many philosophers have not taken the existence of photographic wrongs seriously, beyond the ways photographs can be involved in privacy violations. Part of the reason for this neglect is rooted in the fact that photographs of persons are, by their very nature, representations of persons' bodies or personas, and philosophy has historically not taken persons' embodied natures as seriously as it should. In addition, in our collective social lives, photographs function as natural symbols of the photographic subjects, and philosophy has had surprisingly little to say about symbolic wrongdoing and symbolic valuation more generally. Thus, while my main aim is to make the case for the existence of a heretofore unrecognized type of wrongdoing, I hope this discussion also brings to the fore the variety of ways we can be wronged as embodied social creatures.⁶³

Bryn Mawr College
mcbell@brynmawr.edu

63 I am very grateful to two anonymous referees for the *Journal of Ethics and Social Philosophy* for their insightful comments and criticisms. Thanks also to editors-in-chief Sarah Paul and Matthew Silverstein and managing editor Chico Park for all their help shepherding the essay through the review process. Previous versions of this article were presented at the SIPP@Brown Conference; American Society for Aesthetics (Rocky Mountain Division) Conference; Centre for Ethics, Philosophy, and Public Affairs, University of St. Andrews; and the Center for Human Values, Princeton University. I am very grateful to the audiences at these events for their questions and objections. Liz Harman and Daniel Star provided helpful critical comments on earlier drafts of this essay. Finally, I am grateful to Princeton's Center for Human Values and Bryn Mawr College for support.

REFERENCES

- Allen, Anita. *Unpopular Privacy: What Must We Hide?* Oxford University Press, 2011.
- American Suburb X. "An Interview with Bruce Gilden: Spontaneity, Elegance, and Being Yourself." October 1, 2010. <https://americansuburbx.com/2010/10/interview-interview-with-bruce-gilden.html>.
- Basu, Rima. "The Importance of Forgetting." *Episteme* 19, no. 4 (2022): 471–90.
- Bell, Macalester. "On the Variety of Photographic Wrongs." Unpublished manuscript.
- . "Respecting Photographic Subjects." In *Portraits and Philosophy*, edited by Hans Maes. Routledge, 2019.
- Berstler, Sam. "Bad Question!" *Philosophy and Public Affairs* 51, no. 4 (2023): 413–49.
- Broomberg, Adam, and Oliver Chanarin. *Ghetto*. Trolley, 2003.
- Carnegy-Arbuthnott, Hannah. "Privacy, Publicity, and the Right to Be Forgotten." *Journal of Political Philosophy* 31, no. 4 (2023): 494–516.
- Chadwick, Ruth F. "The Market for Bodily Parts: Kant and Duties to Oneself." *Journal of Applied Philosophy* 6, no. 2 (1989): 129–39.
- Cohen, Jonathan, and Aaron Meskin. "On the Epistemic Value of Photographs." *Journal of Aesthetics and Art Criticism* 62, no. 2 (2004): 197–210.
- Coleman, A. D. "Private Lives, Public Places: Street Photography Ethics." *Journal of Mass Media Ethics* 2, no. 2 (1987): 60–66.
- Danto, Arthur. "The Naked Truth." In *Aesthetics and Ethics: Essays at the Intersection*, edited by Jerrold Levinson. Cambridge University Press, 2001.
- Davies, David. "Susan Sontag, Diane Arbus, and the Ethical Dimensions of Photography." In *Art and Ethical Criticism*, edited by Garry Hagberg. Blackwell, 2008.
- Feinberg, Joel. *Offense to Others*. Oxford University Press, 1984.
- Frowe, Helene, and Jonathan Perry. "Wrongful Observation." *Philosophy and Public Affairs* 47, no. 1, (2019): 104–37.
- Geffer, Philip. "The Theater of the Street, the Subject of the Photograph" *New York Times*, March 15, 2006.
- Gelfert, Axel. "Disattendability, Civil Inattention, and the Epistemology of Privacy." *Philosophical Analysis* 31 (2014): 151–81.
- Goffman, Erving. *Interaction Ritual: Face-to-Face Behavior*. Pantheon Books, 1967.
- . *The Presentation of the Self in Everyday Life*. Doubleday, 1959.
- Hadley, John. "Street Photography Ethics." *Ethical Theory and Moral Practice* 25 (2022): 529–40.

- Herman, Barbara. "Could It Be Worth Thinking About Kant on Sex and Marriage?" In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise M. Antony and Charlotte Witt. Westview Press, 1993.
- . "We Are Not Alone: A Place for Animals in Kant's Ethics." In *Kant on Persons and Agency*, edited by Eric Watkins. Cambridge University Press, 2017.
- Hill, Thomas E., Jr. "Humanity as an End in Itself." *Ethics* 91, no. 1 (1980): 84–99.
- Hopkins, Robert. *Picture, Image, and Experience*. Cambridge University Press, 1998.
- Inness, Julie C. *Privacy, Intimacy, and Isolation*. Oxford University Press, 1992.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge University Press, 1997.
- . *Lectures on Ethics*. Translated by Louis Infield. Edited by Louis White Beck. Hackett Publishing Company, 1963.
- . *The Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge University Press, 1996.
- Korsgaard, Christine. *Fellow Creatures: Our Obligations to the Other Animals*. Oxford University Press, 2018.
- Linfield, Susie. *The Cruel Radiance: Photography and Political Violence*. University of Chicago Press, 2010.
- Marmor, Andrei. "What Is the Right to Privacy?" *Philosophy and Public Affairs* 43, no. 1 (2015): 3–26.
- Martinez, Michel Angela, and Alison Dundes Renteln. "The Human Right to Photograph." In *Images and Human Rights: Local and Global Perspectives*, edited by Nancy Lipkin Stein and Alison Dundes Renteln. Cambridge Scholars Publishing, 2017.
- Nagel, Thomas. "Concealment and Exposure." *Philosophy and Public Affairs* 27, no. 1 (1998): 3–30.
- Olberding, Amy. "Looking Philosophical: Stuff, Stereotypes, and Self-Presentation." *Hypatia* 30, no. 4 (2015): 692–707.
- Rachels, James. "Why Privacy Is Important." *Philosophy and Public Affairs* 4, no. 4 (1975): 323–33.
- Rainsford, Sara. "Ukraine War: The Mothers Going to Get Their Children Back from Russia." *BBC News*, May 30, 2023. <https://www.bbc.com/news/world-europe-65641304>.
- Rini, Regina. "Contingency Inattention: Against Causal Debunking in Ethics." *Philosophical Studies* 177, no. 2 (2019): 369–89.
- . "Deepfakes and the Epistemic Backstop." *Philosophers' Imprint* 20, no. 24 (2020): 1–16.
- Rini, Regina, and Leah Cohen. "Deepfakes, Deep Harms." *Journal of Ethics and*

- Social Philosophy* 22, no. 2 (2022): 143–61.
- Scanlon, Thomas. “Thomson on Privacy.” *Philosophy and Public Affairs* 4, no. 4 (1975): 315–22.
- Sharon, Tamar, and Bert-Jaap Koops. “The Ethics of Inattention: Revitalizing Civil Inattention as a Privacy-Protecting Mechanism in Public Spaces.” *Ethics and Information Technology* 23, no. 3 (2021): 331–43.
- Sontag, Susan. *On Photography*. Picador Press, 1973.
- . *Regarding the Pain of Others*. Farrar, Straus and Giroux, 2003.
- Tadd, G. V. “The Market for Bodily Parts: A Response to Ruth Chadwick.” *Journal of Applied Philosophy* 8, no. 1 (1991): 95–102.
- Thomson, Judith Jarvis. “The Right to Privacy.” *Philosophy and Public Affairs* 4, no. 4 (1975): 295–314.
- Touma, Rafqa. “Melbourne Woman ‘Dehumanized’ by Viral TikTok Filmed Without Her Consent.” *Guardian*, July 14, 2022. <https://www.theguardian.com/technology/2022/jul/14/melbourne-woman-dehumanised-by-viral-tiktok-filmed-without-her-consent>.
- Veliz, Carissa. “Self-Presentation and Privacy Online.” *Journal of Practical Ethics* 9, no. 2 (2022): 30–43.
- Velleman, David. “The Genesis of Shame.” *Philosophy and Public Affairs* 30, no. 1 (2001): 27–52.
- . *How We Get Along*. Cambridge University Press, 2009.
- Walton, Kendall. “Transparent Pictures: On the Nature of Photographic Realism.” *Critical Inquiry* 11, no. 2 (1984): 246–77.
- Warren, Samuel D., and Louis Brandeis. “The Right to Privacy.” *Harvard Law Review* 4, no. 5 (1890): 193–220.

DESIRE SATISFACTION AND TEMPORAL WELL-BEING

TIME FOR A NEW VIEW

Frederick Choo

IT IS COMMON to think that well-being has a temporal dimension—that people can be benefited and be harmed at various times. On most philosophical theories of well-being, accounting for temporal well-being seems straightforward. For example, on the objective list theory, well-being goods consist of objective goods like knowledge, pleasure, and friendship. According to such theories, a person benefits at the times that they have these goods.

However, accounting for temporal well-being is not a clear matter for one prominent theory of well-being—desire satisfactionism.¹ According to desire satisfactionism, a person non-instrumentally benefits if and only if a desire of theirs is satisfied, where a desire is satisfied if and only if the desired object (i.e., the desired state of affairs) obtains.² For many philosophers, desire satisfactionism is a plausible theory of well-being. It captures the idea that what is good *for you* must not alienate you.

To account for temporal well-being, desire satisfactionists need to answer the timing question: *When* does a person benefit from the satisfaction of their desires? Addressing this question is easy in cases involving present-directed desires (desires that are directed towards a present state of affairs).³ Let t_d denote the time a person desires p (where p is some state of affairs); and let t_o denote the time p obtains. In cases where present-directed desires are satisfied, t_d and t_o are the same time. So, a person benefits at that time.

- 1 Although I raise this issue for desire satisfactionism, a similar issue arises for other similar subjectivist theories of well-being (e.g., pro-attitude, value-fulfillment, and aim-achievement theories), as well as for other objective-list theories of well-being that include a subjective component as one of the basic goods.
- 2 Having a desire satisfied is not to be confused with the feeling of satisfaction. If a desired state of affairs occurs, then the relevant desire is satisfied, regardless of whether the person feels satisfied or even knows that the state of affairs occurred.
- 3 In other words, their *content* makes reference to the present time.

However, things are not clear in cases involving past-directed desires (desires that are directed towards a past state of affairs) and in cases involving future-directed desires (desires that are directed towards a future state of affairs), where the person does not have the desire at the time the desired object obtains.⁴ In such cases, since t_d and t_o do not overlap temporally, it is unclear when a person benefits. Henceforth, when I talk about cases involving past-directed desires and cases involving future-directed desires, I am referring to such cases where t_d and t_o do not temporally overlap.

To address such cases, desire satisfactionists have proposed the following views:

Unrestricted Time-of-Desire: A person benefits at t_d .⁵

Time-of-Object: A person benefits at t_o .⁶

Later-Time: A person benefits at whichever time comes later—i.e., a person benefits at t_d in cases involving past-directed desires and benefits at t_o in cases involving future-directed desires.⁷

Fusion: A person benefits at t_{d+o} , which is a fusion of t_d and t_o .⁸

Concurrentism: A person benefits only when t_d and t_o are at the same time.⁹

- 4 There are of course desires that do not fall into any of these categories. We may have desires that are not directed toward any particular time, desires that are directed toward our life as a whole, and desires directed toward eternity. I set these aside, as I think these desires pose a different sort of puzzle and should be addressed differently.
- 5 See Dorsey, "Desire-Satisfaction and Welfare as Temporal" and *A Theory of Prudence*; and Bruckner, "Present Desire Satisfaction and Past Well-Being." Although Bruckner only discusses cases involving future-directed desires, he has informed me that he is also inclined to hold that a person benefits at t_d in cases involving past-directed desires.
- 6 Time-of-object is rarely discussed because it has the implausible implication that someone could be benefited at times that they do not exist—a desired object can obtain before a person is born and after they die. However, see Baber, "Ex Ante Desire and Post Hoc Satisfaction," who defends this view and bites the bullet (64).
- 7 See Lin, "Asymmetrism About Desire Satisfactionism and Time," who proposes *later-time*, though he merely argues that later-time and concurrentism are the most plausible answers to the timing question. To note, Lin calls the view *asymmetrism*, but I have used the name 'later-time' (following Bradley, "Well-Being at a Time") instead since it is more informative and avoids confusion with my view, which also posits an asymmetry.
- 8 See Purves, "Desire Satisfaction, Death, and Time." See also Johansson, who defends Fusion in regard to the badness of death, given that the badness of death must be located in time ("The Time of Death's Badness").
- 9 See Heathwood, "The Problem of Defective Desires"; and Forrester, "Concurrent Awareness Desire Satisfactionism." Bradley also suggests that Concurrentism is the most plausible

Let me elaborate on the latter two views. First, fusion takes t_d and t_o to jointly form the time t_{d+o} . This might seem odd, but Duncan Purves argues that many events take place at fusions of discontinuous times.¹⁰ For example, an academic class may meet only on Mondays and Fridays, so the class occurs at a fusion of the time t_{M+F} . Another plausible example is a soccer match, which can be said to occur at a fusion of the two forty-five-minute halves, even though there is a break between the halves.

Second, fusion is not the view that a person benefits at both t_d and t_o . Proponents of fusion hold that there are no times *within the fusion* at which you benefit.¹¹ So on their view, a person benefits at t_{d+o} , even though a person benefits neither at t_d nor at t_o .¹²

Third, concurrentism differs from the other views in that it rejects that people can benefit in cases involving past-directed or future-directed desires. As noted above, according to concurrentism, a person can benefit only when t_d and t_o are at the same time. This reveals that one way we might answer the timing question is by questioning the assumption that we benefit in the first place in the two cases under discussion instead of trying to specify a time when we benefit in such cases.

In this article, I advance a new view that I call *no-future time-of-desire*.¹³

No-Future Time-of-Desire: (a) In cases involving past-directed (and present-directed) desires, a person benefits at t_d , and (b) in cases involving future-directed desires, a person cannot benefit.

view for desire satisfactionists ("Well-Being at a Time"). To note, in a later paper, Heathwood defends *subjective desire satisfactionism*, which holds that well-being consists of *believing* that one is getting (will get or has gotten) what one desires ("Desire Satisfactionism and Hedonism"). Forrester also defends a similar view ("Concurrent Awareness Desire Satisfactionism"). I think this epistemic requirement is what leads them towards concurrentism. I set both their versions of subjective desire satisfactionism aside since this diverges from standard desire satisfactionism, and I think their views should be handled very differently since there is now another distinct time to consider—the time one has the belief.

10 Purves, "Desire Satisfaction, Death, and Time," 806.

11 Johansson, "The Time of Death's Badness," 475; and Purves, "Desire Satisfaction, Death, and Time," 805.

12 I find this view somewhat puzzling, especially given Purves's examples ("Desire Satisfaction, Death, and Time," 806). In the case of the academic class, if a class occurs at a fusion of time t_{M+F} , it is because part of the class occurs on t_M and part of the class occurs on t_F . So, if someone benefits at t_{d+o} , it seems that part of the benefit should be incurred at t_d and part of the benefit should be incurred at t_o .

13 Although Bradley mentions the view ("Well-Being at a Time," 10), he neither endorses nor defends it. I worked on this view earlier and independently, and this article is the first defense of this view at length.

Although this view may seem odd, my goal is to argue that no-future time-of-desire is superior to the other existing views in the literature.

For easy reference, table 1 summarizes the implications of the views above with respect to past-directed and future-directed desires.

Table 1. *Different Views*

View	Past-directed desire (t_o precedes t_d)	Future-directed desire (t_d precedes t_o)
Unrestricted time-of-desire	Benefit at t_d	Benefit t_d
Time-of-object	Benefit at t_o	Benefit at t_o
Later-time	Benefit at t_d	Benefit at t_o
Fusion	Benefit at t_{d+o}	Benefit at t_{d+o}
No-future time-of-desire	Benefit at t_d	No benefit
Concurrentism	No benefit	No benefit

Here is the structure of this article. Section 1 advances the first argument for no-future time-of-desire using a pair of cases—one involving past-directed desires and the other involving future-directed desires. I argue that only no-future time-of-desire accommodates our intuitions in both cases. The second argument for no-future time-of-desire appeals to two plausible principles, which I defend further. The first principle is Dale Dorsey's *synchronic resonance constraint* (defended in section 2).¹⁴ The second principle is Eden Lin's *all-conditions-met principle* (defended in section 3).¹⁵ I argue that only no-future time-of-desire and concurrentism can accommodate both principles. In section 4, I then argue that no-future time-of-desire is preferable to concurrentism. Sections 5 through 9 handle various objections. Section 10 concludes.

Before moving on, I must note that there is one more view that I have set aside for this article: *atemporalism*. Atemporalism says that in all cases where people benefit, people benefit atemporally.¹⁶ Harry Silverstein, for example, thinks that it does not make sense to ask *when* something is good for the

14 Dorsey, "Desire-Satisfaction and Welfare as Temporal," 156–57. The naming of this principle comes from Lin, "Asymmetrism About Desire Satisfactionism and Time," 179.

15 Lin, "Asymmetrism about Desire Satisfactionism and Time," 169. Lin calls this the "first principle." I have changed the name to make it more informative.

16 Silverstein, "The Time of the Evil of Death"; and Bramble, *The Passing of Temporal Well-Being*. See also Johansson, who defends atemporalism against an objection regarding the badness of death ("The Time of Death's Badness"), and Bykvist, who suggests atemporalism ("Sumner on Desires and Well-Being," 481).

person.¹⁷ This applies no matter how a person's desires and the desired object are located temporally. So, atemporalism rules out temporal well-being entirely.

Holding that there is no temporal well-being at all is deeply unintuitive. We commonly speak in ways that imply that things can be good for us at particular times. For example, we often ask people how well their day or week was. We also think that people have different levels of well-being over time. For example, I might say that although things were not going well for me *when* I first started my own business, things are going well for me *now*. The idea that there is temporal well-being features not just in ordinary talk but in multiple fields of academic studies. Social scientists claim to measure well-being at times and to track well-being changes over time. Philosophers not only talk about temporal well-being, but they also discuss the shape of a person's life over time.¹⁸ Those working in public policy often work at creating policies to improve the well-being of people. So, atemporalism has difficulty accounting for our ordinary thoughts about well-being. It would be surprising if we were totally wrongheaded in thinking about well-being. Hence, I set aside atemporalism. My goal is to argue with those who agree that at least some benefits can be located in time.

1. TWO CASES

In this section, I advance my first argument for no-future time-of-desire by appealing to two cases.¹⁹ First, consider a case involving past-directed desires.

Singing: Suppose that when Helen is at university (Youthful-Helen), she loves singing karaoke with her friends. Although she sings well, she does not care about whether she sings well—she just likes singing. A few years later, she loses interest, gets busy, and no longer goes to sing karaoke. Now suppose when Helen is fifty years old (Older-Helen), she picks up interest in singing again. She is very serious about it this time and cares about singing *well*. Older-Helen looks back at pictures

17 Silverstein, "The Time of the Evil of Death."

18 As Campbell notes, "It seems better to have a life that begins poorly and ends well than a life that begins well and ends poorly" ("When the Shape of a Life Matters," 565).

19 To note, although some cases like these have been raised by other authors, these authors typically do not consider both cases involving past-directed desires and cases involving future-directed desires. Furthermore, some of their cases seem incompatible with certain restrictions that desire satisfactionists may want to hold (e.g., counting only self-regarding desires or idealized/informed desires). My cases involve self-regarding desires, and one can read the cases as rational adults forming their desires under idealized conditions. I discuss this more in the conclusion.

of Youthful-Helen at karaoke. Helen does not know how well Youthful-Helen sang, but she desires that she sang well all those past times.²⁰

Intuitively, it seems good for Helen if Youthful-Helen really did sing well. When does Helen benefit from having sung well? Intuitively, Helen benefits when she has the desire to sing well. This intuition is shared by several philosophers who write about desire satisfactionism.²¹ It seems implausible for desire satisfactionists to say that Helen benefitted in the past—as time-of-object claims—since Youthful-Helen was not concerned about whether she sang well then. She cannot benefit in the past from a satisfied desire, since she had no desire then.

Some might be hesitant to hold that Helen benefits, since Older-Helen does not know whether Youthful-Helen sang well. To address this, it is important to note that desire satisfactionists standardly hold that a state of affairs can affect a person's well-being even if they never know about it, and it never affects their experiences.²² Various cases illustrate this, the most obvious examples involving harm. Suppose you desire that your privacy is not violated by people spying on you and that your friends do not make fun of you behind your back. Intuitively, violations of your privacy and betrayals of friendship seem bad for you, even if you never come to know about them. There are also examples involving benefits. For example, if you desire that people listen to your music and love it, then it seems intuitive to think that it is good for you that people listen to your music and love it, even if you never come to know about this. Desire satisfactionists hold that it is a virtue of their view that they can accommodate such cases where a person's well-being is affected, even though the person does not know whether the desired state of affairs has occurred, and it never affects their experiences. So, if we grant this claim to desire satisfactionists, it seems right to hold that Helen benefits despite not knowing that she got what she wanted.

One might object that the intuition that Helen benefits from having sung well in the past is because singing well would have led to the satisfaction of other desires in the past.²³ Thus, the intuition is really about the instrumental benefits of having sung well. To address this, we can imagine that Youthful-Helen did sing

20 We might add that Older-Helen is not anxious or worried about not having sung well.

21 See, for example, Sarch, "Desire Satisfactionism and Time," 232–33; Lin, "Asymmetrism About Desire Satisfactionism and Time," 167; and Purves, "Desire Satisfaction, Death, and Time," 803.

22 Exceptions to this are Heathwood, "Desire Satisfactionism and Hedonism"; and Forrester, "Concurrent Awareness Desire Satisfactionism." Both their views require the person to believe that they are getting what they want. As mentioned in note 9 above, I set aside their views, as they diverge from standard desire satisfactionism and should be handled very differently.

23 I thank an anonymous reviewer for this objection.

well in the past, and consequently, any instrumental benefits she gained from having sung well has manifested in her life. Now, we can ask: Is there a difference in well-being when comparing a possible world in which Older-Helen has the desire to have sung well and another possible world in which Older-Helen has no such desire? The answer seems yes: Helen is better-off in the first world, where she had a desire to have sung well satisfied. Hence, the intuition that Helen benefits is not due to the instrumental goods of having sung well.

Which views can accommodate the intuitions above? Out of the different views, only no-future time-of-desire, unrestricted time-of-desire, and later-time accommodate these intuitions since these views locate the benefit at t_d in cases involving past-directed desires. Time-of-object gives the implausible result that Helen benefitted in the past at t_o . Concurrentism gives the implausible result that Helen does not benefit at all. Fusion gives the verdict that Helen benefits only at a fusion of the times Youthful-Helen sings well and the times that Older-Helen desires to have sung well in the past. Although this result is not implausible (as in the case of time-of-object and concurrentism), it is not a natural one. It is not intuitive to judge that Helen benefits at the fusion of those times (and does not benefit at either of those times). So, fusion is less favorable as it lacks the intuitive pull, though I think the Singing case does not rule it out.

Next, let us turn to a second case involving future-directed desires.

Living in Japan: Suppose that when Gary is twenty-one years old (21-Gary), he loves Japanese culture. He desires to live in Japan in the future. He joins a Japanese company hoping they will send him to Japan. However, by his late twenties, Gary has a change of heart. He is no longer interested in Japanese culture and has completely lost the desire to live in Japan.²⁴ When Gary is thirty years old (30-Gary), his company requires him to be posted to Japan for two years. So, he ends up living in Japan.

Intuitively, it seems that living in Japan per se did not increase Gary's well-being.²⁵ Again, the intuition that a person does not benefit in cases involving

24 Gary should not be taken as having some con attitude towards living in Japan (e.g., desiring not to live in Japan). Rather, Gary is simply neutral towards it. He neither has a desire to live in Japan nor a desire to not live in Japan.

25 One might object that the intuition is because living in Japan would frustrate Gary's other desires. However, we can imagine that going to Japan would not frustrate his other desires—he would be able to satisfy his other desires regardless of whether he is living in Japan or staying where he is.

future-directed desires is shared by several philosophers who write about desire satisfactionism.²⁶

Time-of-object and later-time give the wrong verdict since they entail that Gary benefits at thirty years old. This seems implausible since 30-Gary no longer desires to live in Japan. It does not seem good for Gary to get something when he no longer desires it. Unrestricted time-of-desire also gives the wrong verdict since it entails that Gary benefits at twenty-one years old. This seems implausible since 21-Gary's desire is still unsatisfied then—it is years before he gets what he wants. Hence, it seems implausible to say that Gary benefited at twenty-one years old.²⁷ Fusion gives the verdict that Gary benefits at the fusion of twenty-one and thirty years old (and does not benefit at either of those times). While it does not give the implausible result that Gary benefits at twenty-one years old or the implausible result that Gary benefits at thirty years old, fusion still clashes with the intuition that Gary does not benefit at all.²⁸

Out of the different views, only no-future time-of-desire and concurrentism accommodate the intuition that Gary did not benefit. Both of these theories hold that a person does not benefit in cases involving future-directed desires. All the other theories give the wrong verdict as they entail that Gary benefits.

Since only no-future time-of-desire accommodates the intuitions in both *Singing* and *Living in Japan*, no-future time-of-desire enjoys intuitive support over the other competing views.

At this point, some may object that cases involving future-directed desires are not so straightforward because future-directed desires seem conditional on their own persistence.²⁹ It would be odd if 21-Gary desires to live in Japan in the future even if he would no longer want to in the future. Plausibly, 21-Gary's

26 See, for example, Sumner, *Welfare, Happiness, and Ethics*, 126–27; Bykvist, “The Moral Relevance of Past Preferences”; Heathwood, “The Problem of Defective Desires,” 490; Portmore, “Desire Fulfillment and Posthumous Harm,” 30; Bradley, *Well-Being and Death*, 27, and “Well-Being at a Time,” 7; and Bruckner, “Present Desire Satisfaction and Past Well-Being.”

27 In addition, Portmore (in “Desire Fulfillment and Posthumous Harm”) and King (in “Pulling Apart Well-Being at a Time and the Goodness of a Life” and “The Good of Today Depends Not on the Good of Tomorrow”) have argued that holding that past (and present) well-being depends on future events results in various other counterintuitive results. King also argues the idea that present well-being may depend on future events conflicts with the assumptions and practices in empirical research on well-being (“The Good of Today Depends Not on the Good of Tomorrow,” 2371, 2378–79).

28 Even if we hold that Gary benefits, it is not intuitive to judge that Gary benefits at the fusion of twenty-one and thirty years old (and does not benefit at either of those times). So, fusion lacks the intuitive pull again.

29 I thank multiple philosophers who have raised this worry.

desire to live in Japan is a conditional desire—21-Gary wants to live in Japan in the future *only if* his future self still has such a desire when it occurs. If 21-Gary's desire is conditional, then it is not satisfied since 30-Gary lacks the desire. So, all the different views entail that Gary does not benefit in Living in Japan. This makes later-time, unrestricted time-of-desire, and no-future time-of-desire on par in capturing the intuitions of both Singing and Living in Japan.

To address this objection, we must first ask whether *all* future-directed desires *must* be conditional: either all future-directed desires must be conditional, or they can be unconditional sometimes. Whichever way we go, I think we still get support for no-future time-of-desire.

Consider the latter view (which I lean towards). Here is a plausible example of an unconditional future-directed desire. Given my love for my wife, I desire that I still take care of my wife when I am old, regardless of whether I desire so then.³⁰ Suppose that it is possible for future-directed desires to be unconditional in this way. In that case, we can imagine 21-Gary as having an *unconditional* desire in Living in Japan that is satisfied when he is thirty years old. Even given this specification, it still seems that Gary does not benefit. So, my appeal to Living in Japan can proceed as before.

Now suppose instead that all future-directed desires must be conditional. Notice then that such desires can never be satisfied in *all* cases involving future-directed desires. This is because a person lacks the desire at t_0 in such cases, and hence, the condition cannot be met. The result of this view is that a person cannot benefit in cases involving future-directed desires. This again gives no-future time-of-desire an advantage over competing views since no-future time-of-desire claims exactly this—a person cannot benefit in cases involving future-directed desires.

To end this section, I must note that the intuitive support for no-future time-of-desire is not universal. Out of the people I have shared these cases with, it seems that the intuitions are prevalent among a significant number of people (at least, if they were sympathetic to desire satisfactionism or if they included satisfied desires on their objective list). But not everyone has the same intuitions regarding both Singing and Living in Japan. Also, philosophers who write about desire satisfactionism typically express their intuitions either only when considering cases involving past-directed desires or only when considering cases involving future-directed cases. So, it may not be clear that they have the intuitions I have expressed regarding *both* types of cases.

30 A reviewer provides another example of unconditional future-directed desires: "I hope that when I am older, I still desire peace and equality." However, this example only establishes that there are unconditional *second-order* future-directed desires. What is of concern here, however, are first-order desires.

Therefore, I want to advance a second argument for no-future time-of-desire by defending two principles in the next two sections. These principles jointly provide theoretical backing for no-future time-of-desire.

2. THE SYNCHRONIC RESONANCE CONSTRAINT

A major motivation for desire satisfactionism is the resonance constraint, which states that in order for something to be good for you, it must resonate with you.³¹ The idea is that for something to be good *for you*, you must have a favorable attitude toward it. As Peter Railton says, “What is intrinsically valuable for a person must have a connection with what [the person finds] in some degree compelling or attractive, at least if he were rational and aware. It would be an intolerably alienated conception of someone’s good to imagine that it might fail in any such way to engage him.”³² Desire satisfactionism accommodates the resonance constraint by holding that something is good for a person only if they desire it. When a person desires a state of affairs, that state of affairs resonates with the person.

In arguing for unrestricted time-of-desire, Dorsey proposes that desire satisfactionists further accept the synchronic resonance constraint.

Synchronic Resonance Constraint: A person does not benefit at time t unless they have the relevant desire at t .³³

Given that desire satisfactionists accept the resonance constraint because they think well-being goods should not alienate a person, it is natural to think that a person can benefit only at the time the object resonates with them. Without the synchronic resonance constraint, a person could benefit at times when the desired object does not resonate with them. As Dorsey says, if one does not accept the synchronic resonance constraint, then this “would leave it open that I can lack a desire for ϕ at a particular time, but that ϕ is nevertheless

31 This requirement is also often referred to as the nonalienation requirement and internalism. See, for example, Rosati, “Internalism and the Good for a Person,” 297–326; Railton, “Facts and Values,” 47–48; and Dorsey, “Why Should Welfare ‘Fit’?” 685–708.

32 Railton, “Facts and Values,” 47. Railton himself includes “at least if he were rational and aware” because he holds to an idealized version of desire satisfactionism. For nonidealized versions, what matters is whether an object resonates with the actual person (rather than with the counterfactual idealized person).

33 Dorsey, “Desire-Satisfaction and Welfare as Temporal,” 156–57, and *A Theory of Prudence*, 193. See Bradley (“Well-Being at a Time,” 4) and Purves (“Desire Satisfaction, Death, and Time,” 810), who also think desire satisfactionists must hold to the synchronic resonance constraint.

good for me at that time.”³⁴ Rejecting the synchronic resonance constraint means that a person could be benefited at times when they are alienated from a state of affairs. Hence, desire satisfactionists should accept the synchronic resonance constraint.

Given the synchronic resonance constraint, which views can accommodate this principle? No-future time-of-desire, unrestricted time-of-desire, and concurrentism can accommodate it since these views always locate benefits at t_d . Time-of-object and later-time violate this constraint because in cases involving future-directed desires, they allow a person to benefit at a time they lack the relevant desire.

For fusion, Purves argues that it is a virtue of fusion that it can accommodate the synchronic resonance constraint. He says, “Any fusion of times at which the object of a desire benefits a person is one at which both the desire and its object obtain, as required by the synchronic resonance constraint.”³⁵ I find this objectionable. Fusion says that Gary benefits at t_{d+o} , a fusion of t_d and t_o . If Gary benefits at t_{d+o} , the synchronic resonance constraint requires that Gary has the desire at t_{d+o} . However, Gary has the desire only at t_d , which is *part* of t_{d+o} . He does not have the desire at t_{d+o} . Hence, fusion fails to accommodate the synchronic resonance constraint.

Fusion proponents might try to avoid my objection by holding that if a person desires p at t_d , then the person desires p at t_{d+o} . This is because a fusion of time may take on the properties of its temporal parts. However, this view faces two problems. First, it is doubtful that desiring p at t_d entails that the person desires p at t_{d+o} . This seems deeply unintuitive. This is most evident when we consider other mental states. For example, if I enjoyed only the first half of a soccer match, it does not seem that I enjoyed myself at the fusion of the first half and second half of the soccer match. Another example: if I believe p during class on Monday but no longer believe p during class on Friday, it does not seem that I believe p at the fusion of time t_{M+F} . Fusion proponents bear the burden of showing that desiring p at t_d entails that the person desires p at t_{d+o} .

Second, I think such a view results in a contradiction. Suppose that at t_d , the person desires p ; and at t_o , the person lacks a desire for p . If we apply the view of fusion proponents, we get the result that the person both desires p and does not desire p at t_{d+o} . We would have to say that p both resonates and does not resonate with the person at the same time. This seems like a contradiction.

34 Dorsey, “Desire-Satisfaction and Welfare as Temporal,” 156.

35 Purves, “Desire Satisfaction, Death, and Time,” 810.

Let me now consider one objection to the synchronic resonance constraint from Lin.³⁶ Lin asks us to consider a similar principle, the *synchronic resonance constraint*_o:

*Synchronic Resonance Constraint*_o: A person does not benefit at any time from a particular event unless they have a desire toward that event at the time which the event occurs.³⁷

If the synchronic resonance constraint_o is true, then concurrentism is true. People can benefit only when t_d and t_o are at the same time. Lin points out that unrestricted time-of-desire proponents reject the synchronic resonance constraint_o even though they hold to the resonance constraint. Since desire satisfactionists are motivated by the resonance constraint but need not accept the synchronic resonance constraint_o, Lin thinks that desire satisfactionists need not accept the synchronic resonance constraint as well.

I think Lin's objection fails. The resonance constraint motivates the synchronic resonance constraint but does not similarly motivate the synchronic resonance constraint_o. Since the resonance constraint is concerned that well-being goods do not alienate a person, desire satisfactionists should think that a person cannot benefit at a time when a purported good does not resonate with them. For if we suppose that a person can benefit at a time a purported good does not resonate with them, this would mean that an object could be good for a person at a time when the object alienates them. This seems to run against the spirit of the resonance constraint, which does not want well-being goods to alienate a person. Surely, those moved by the resonance constraint will find it hard to allow a person to benefit at times when they are alienated. What seems to flow naturally from the resonance constraint is that a person can benefit only at times when the desired object resonates with them. So, if desire satisfactionists are concerned that well-being goods must resonate with a person, desire satisfactionists should think that a person benefits at the time they have the desire rather than at a time they lack the desire. Therefore, those who accept the resonance constraint have reason to accept the synchronic resonance constraint. In contrast, there is no similar argument from the resonance constraint to the synchronic resonance constraint_o. It is unclear why the resonance constraint would give us reason to think that a person can benefit only if they desire the desired object *when it obtains*. Therefore, what motivates the synchronic resonance constraint does not similarly motivate the synchronic resonance constraint_o.

36 Lin, "Asymmetrism About Desire Satisfactionism and Time," 179.

37 Lin, "Asymmetrism About Desire Satisfactionism and Time," 179.

One might object by arguing that a desired object can resonate with a person only if they desire the object *at the time it obtains*. Given this, the resonance constraint similarly motivates the synchronic resonance constraint.₃₈ In reply, I think it is unintuitive to think that a desired object can resonate with a person only if they desire the object *while it obtains*. Consider Living in Japan again. Since desire satisfactionists understand resonance in terms of desiring, desire satisfactionists should hold that living in Japan resonates with 21-Gary because 21-Gary has the relevant desires then. Therefore, desire satisfactionists should reject the idea that a desired object can resonate with people only if they desire the object *while it obtains*. Future states of affairs can resonate with people presently because people can presently desire future states of affairs.³⁸ So the resonance constraint does not motivate the synchronic resonance constraint.₃₉

3. THE ALL-CONDITIONS-MET PRINCIPLE

Now let me turn to a second principle. Lin argues for (what I call) the *all-conditions-met principle*.

All-Conditions-Met Principle: “You do not receive a particular benefit at *t* unless, at *t*, all of the necessary conditions [for] receiving that benefit have been met.”³⁹

Lin says that a “condition has been met at *t* just if it either *is* met at *t* or *was* met at some time prior to *t*.”⁴⁰ The all-conditions-met principle prevents a person from benefiting until the earliest time that all the necessary conditions for receiving the benefit have been met.

The all-conditions-met principle seems very plausible. It is intuitive to think that something cannot obtain until all the necessary conditions for the thing obtaining are met first. For example, even if I have submitted my tenure application and it will be approved next month, I am not tenured presently because the necessary conditions for being tenured are not met yet—the application has not yet been approved. Similarly, all necessary conditions for benefitting must be met first before a person can benefit at a time.

For desire satisfactionists, the necessary conditions for benefitting are *desiring p* and *p obtains*. So, the all-conditions-met principle prevents a person from

38 The same applies to past states of affairs as seen in Singing. People can presently desire a past state of affairs, and so past states of affairs can resonate with people presently.

39 Lin, “Asymmetrism About Desire Satisfactionism and Time,” 169–71. Lin calls this the First Principle. See also Purves, who finds this principle attractive (“Desire Satisfaction, Death, and Time,” 809).

40 Lin, “Asymmetrism About Desire Satisfactionism and Time,” 169.

benefiting at a time when they have not yet gotten what they wanted, and it prevents a person from benefiting at a time when they have not yet started having the relevant desire.

One objection to the all-conditions-met principle comes from Dorsey.⁴¹ He argues that the all-conditions-met principle implies that we would have to assign “prudential benefits to times at which those benefits alienate” in cases involving future-directed desires.⁴² We would have to hold that 30-Gary benefits since both conditions are met then, even though 30-Gary does not have the relevant desire anymore. This would violate the synchronic resonance constraint. So, we should reject the all-conditions-met principle.

However, I think Dorsey is mistaken. The all-conditions-met principle states only a necessary condition for benefitting at a time, not a sufficient condition. So, it does not entail that 30-Gary benefits. One can hold to the all-conditions-met principle and yet maintain that 30-Gary does not benefit (as no-future time-of-desire claims).

Given the all-conditions-met principle, which views can accommodate it? Neither time-of-object nor unrestricted time-of-desire can accommodate the all-conditions-met principle. In cases involving past-directed desires, time-of-object says that you benefit at t_o , although you do not have the relevant desire yet. In cases involving future-directed desires, unrestricted time-of-desire says that you benefit at t_d , although the desired object has not yet been obtained. So, both views allow a person to benefit at a time when only one of the necessary conditions for receiving that benefit has been met.

All the other views seem to be able to accommodate the all-conditions-met principle. Concurrentism allows us to benefit only in cases involving present-directed desires, and in such cases, all the necessary conditions are met by the time of benefit because both conditions obtain at the time of benefit. For fusion, Purves argues that fusion can accommodate the all-conditions-met principle because the “necessary conditions have been met at the fusion because they both obtain at the fusion.”⁴³ For later-time, Lin argues that later-time accommo-

41 Dorsey, *A Theory of Prudence*, 199–200.

42 Dorsey, *A Theory of Prudence*, 199.

43 Purves, “Desire Satisfaction, Death, and Time,” 809. One reviewer objects that it is not clear that fusion satisfies the all-conditions-met principle because it is not clear *when* t_{d+o} occurs. Does it occur before/after t_d or before/after t_o ? I think fusion proponents would say that the fusion of time does not occur before or after any of its parts. Rather, the time is its parts constituted together. For example, the academic class that meets on Monday and Friday takes place at t_{M+F} . Yet the time the class meets is neither before/after t_M nor before/after t_F . In any case, it is not crucial to this article whether fusion satisfies the all-conditions-met principle. If fusion fails to accommodate the all-conditions-met principle, then all the better for my argument.

dates the all-conditions-met principle.⁴⁴ This is because in both cases involving past-directed and cases involving future-directed desires, a person benefits at a time when both conditions have been met.

Finally, consider no-future time-of-desire. In cases involving past-directed desires, a person benefits at t_d , and the desired object has obtained in the past. So, both necessary conditions have been met. In cases involving future-directed desires, a person does not benefit. Hence, no-future time-of-desire does not run into the same problem as unrestricted time-of-desire.

Let us take stock. I have defended the synchronic resonance constraint and the all-conditions-met principle. I have argued that no-future time-of-desire accommodates both principles. In contrast, I have argued that almost all of the other competing views cannot satisfy one or both of the principles, as shown in table 2.

Table 2. The Views and Two Principles

View	Synchronic resonance constraint	All-conditions-met principle
Unrestricted time-of-desire	Yes	No
Time-of-object	No	No
Later-time	No	Yes
Fusion	No	Yes
No-future time-of-desire	Yes	Yes
Concurrentism	Yes	Yes

Although I have argued that no-future time-of-desire can accommodate both principles, concurrentism can also accommodate both principles.⁴⁵ In the next section, I offer two reasons to prefer no-future time-of-desire over concurrentism.

4. CONCURRENTISM OR NO-FUTURE TIME-OF-DESIRE?

There are two reasons why I think no-future time-of-desire is preferable to concurrentism. First, we can look back to cases involving past-directed desires that

⁴⁴ Lin, "Asymmetrism About Desire Satisfactionism and Time," 169.

⁴⁵ To note, Bradley argues that desire satisfactionists should further accept the synchronic resonance constraint, and hence hold to concurrentism ("Well-Being at a Time," 7). He argues for this by pointing out that we intuitively judge that a person does not benefit in cases involving future-directed desires (such as in *Living in Japan*). I think this argument fails because both concurrentism and no-future time-of-desire can accommodate such intuitions. We need a further argument to see which view is preferable.

intuitively seem to affect a person's well-being.⁴⁶ Recall the Singing case from section 1. If Helen desired to have sung well in the past, it seems good for her that Youthful-Helen really did sing well. No-future time-of-desire can accommodate our intuitions in such cases. However, concurrentism cannot account for how such past-directed desires can seem to affect a person's well-being if they are satisfied. So, we have good reason to reject concurrentism.⁴⁷

Second, various philosophers argue that since desire satisfactionists accept spatial distance in their theory, they should also allow temporal distance.⁴⁸ Ben Bradley, Dorsey, and Purves each note that desire satisfactionists hold that states of affairs can affect our well-being without making a difference in our experiences. As mentioned in section 1, this seems intuitive when considering various cases (e.g., a violation of privacy, a betrayal of friendship). Based on this idea of spatial distance, Purves (and the others) criticizes concurrentism: "if [we] accept that spatial distance between the experience of the desirer and the object of her desire is irrelevant to whether her desire counts as satisfied, then why should temporal distance matter?"⁴⁹

While I think the general thought is intuitive, I find it odd to understand *spatial* distance in terms of whether an event affects a person's experiences (or enters their knowledge). I think this sense of spatial distance is not the right thing to compare with temporal distance.

Spatial distance is more naturally understood in terms of whether an event occurs at the same physical location as the person who benefits. With this understanding of spatial distance, we can still advance a similar argument (call this the *physical-temporal location argument*). Suppose I am in Singapore, while my wife is an astronaut on the moon. I desire that she does not cheat on me while she is there. Even though we are in different physical locations, desire satisfactionists agree that it is bad for me if she cheats on me. Since desire satisfactionists hold that spatial distance (in terms of the event occurring at a different physical location from the person who benefits) does not matter to whether people can benefit, desire satisfactionists should similarly hold that temporal distance does not matter to whether people can benefit.

46 See Purves, "Desire Satisfaction, Death, and Time," 803; and Dorsey, "Desire-Satisfaction and Welfare as Temporal," 157–58.

47 For the same reason, we should reject the synchronic resonance constraint.

48 See Bradley, *Well-Being and Death*, 23; Dorsey, "Desire-Satisfaction and Welfare as Temporal," 157–58, and *A Theory of Prudence*, 194–95; and Purves, "Desire Satisfaction, Death, and Time," 803. See also Fischer, who makes such an argument, though without reference to desire satisfactionism in particular (*Our Stories*, 45–46). In a later paper, however, Bradley abandons this argument ("Well-Being at a Time," 9).

49 Purves, "Desire Satisfaction, Death, and Time," 803.

Given this, concurrentism is problematic. It holds that people cannot benefit in cases involving past-directed desires and cases involving future-directed desires because the desire and the desired object are temporally distant in such cases. In contrast, no-future time-of-desire allows people to benefit despite temporal distance.⁵⁰ Hence, no-future time-of-desire is preferable to concurrentism.

Bradley rejects the physical-temporal location argument.⁵¹ He points out that although we have a notion of temporal well-being, we do not have a notion of spatial well-being. We do not make statements like “I am benefited here.”⁵² This explains why spatial distance does not matter. However, since we have a notion of temporal well-being, he maintains that temporal distance plausibly matters.

In response, I do not see why whether we have a notion of spatial well-being would matter. We need to keep the following two questions separate.

Location Question: Where does a person benefit?

Benefit_s Question: Can a person benefit if their desire and the object are spatially distant?

We might think that the location question does not make sense if there is no notion of spatial well-being. But notice that we can (and do) still give an affirmative answer to the benefit_s question because these are distinct questions. Similarly, we need to keep the following two questions separate.

Timing Question: When does a person benefit?

Benefit_t Question: Can a person benefit if their desire and the object are temporally distant?

We might propose various answers to the timing question. But it is unclear why thinking that there is an answer to the timing question (i.e., the mere fact that

50 One reviewer objects that no-future time-of-desire is problematic as it still restricts temporal distance partially since it proposes that one cannot benefit in cases involving future-directed desires. In reply, my view does not hold that we cannot benefit in such cases *because the desire and the desired object are temporally distant*. Instead, the reason why we do not benefit in such cases is because of the all-conditions-met principle and my argument to come in section 6, not because the desire and the desired object are temporally distant.

51 Bradley, “Well-Being at a Time,” 9.

52 I am slightly worried about what we should make of the fact that we do not make such statements. Although we typically do not talk about spatial well-being, it seems natural to say that we benefit wherever we are located.

we have a notion of temporal well-being) should entail that we answer no to the benefit_t question.

How the physical-temporal location argument should be understood is this: since desire satisfactionists say yes to the benefit_s question, it seems that they should say yes to the benefit_t question as well. Whether or not desire satisfactionists think that the location question and the timing question have answers is irrelevant to how we should answer the benefit_s and benefit_t questions, respectively.

I conclude then that no-future time-of-desire is superior to the other existing views in the literature. Let me now turn to some objections.

5. UNRESTRICTED TIME-OF-DESIRE AND THE ALL-CONDITIONS-MET PRINCIPLE

Earlier, I argued that although unrestricted time-of-desire satisfies the synchronic resonance constraint, it violates the all-conditions-met principle. However, one might object that unrestricted time-of-desire can actually satisfy the all-conditions-met principle. In defending unrestricted time-of-desire, Donald Bruckner argues that truths about the future can make it the case that a person's desire is satisfied in the present, and so they can benefit prior to the state of affairs occurring.⁵³ For example, suppose that in 2025 (now), Eve desires to be pregnant in 2026; and she will get pregnant in 2026. So, it is true in 2025 (now) that she is pregnant in 2026. Since it is true now that she is pregnant in 2026 and she presently desires to be pregnant then, her desire is satisfied presently in 2025.

Bruckner argues that this is plausible because there are cases where future states of affairs seem to determine the status of past states of affairs. For example, "A shooting acquires the status of a killing only if the victim dies as a result of the gunshot, which may be some time later."⁵⁴ Another example is the truth value of propositions that reference the future. If a person says that there will be a sea battle tomorrow, the truth value of the statement uttered now seems to depend on what happens tomorrow.⁵⁵ Similarly, although the desired object obtains in the future, it is plausible to think that Eve's desire is satisfied in the present, and so Eve benefits presently because the desired state of affairs will obtain. Therefore, unrestricted time-of-desire can accommodate the all-conditions-met

53 Bruckner, "Present Desire Satisfaction and Past Well-Being," 18, 25–26. See also Bradley, *Well-Being and Death*, 23–24.

54 Bruckner, "Present Desire Satisfaction and Past Well-Being," 25–26.

55 Bruckner, "Present Desire Satisfaction and Past Well-Being," 18.

principle by holding that the relevant necessary conditions for benefitting are actually *desiring p* and *it is true that p obtains at some time*.

I think this argument is unsuccessful for three reasons. First, intuitively, the two necessary conditions for having a desire satisfied are actually (a) *desiring p* and (b) *p* obtains. Even if it is true now in 2025 that Eve gets pregnant in 2026, it seems that her desire is not yet satisfied in 2025 because she has not gotten what she wants in 2025. Instead, her desire will be satisfied in 2026 when she gets what she wants. The 'it is true that' talk is simply irrelevant to having a desired satisfied at a certain time.⁵⁶

Second, I think we should not hold that the necessary conditions for benefitting include a 'it is true that' clause. Suppose a desire satisfactionist proposes that we add the 'it is true that' clause to *both* necessary conditions. The necessary conditions for having a desire satisfied would then be (a) it is true that the person desires *p* at some time, and (b) it is true that *p* obtains at some time. A desire would count as satisfied at the time at which it is true that a person has the desire and at which it is true that the object obtains. For example, if, in 1999, it is true that Eve will form the desire in 2025 to be pregnant in 2026, and if, in 1999, it is true that Eve will be pregnant in 2026, then we would get the result that Eve's desire was satisfied in 1999. This result is absurd. How can Eve's desire be satisfied in 1999, years before she has the desire and years before she gets pregnant? From this example, it is clear that whether *it is true* that a person has the desire at a future time is irrelevant to counting a desire as satisfied. Similarly, we should think that whether *it is true* that the event obtains at a future time is irrelevant to counting a desire as satisfied. Those who think that the 'it is true that' clause matters to one necessary condition but not to the other bear the burden of offering some explanation for this asymmetry. Without such an explanation, we should reject adding the 'it is true that' clause to any of the necessary conditions.⁵⁷

Third, in many cases, we do not just desire that a proposition is true. Rather, we desire the obtaining of the future state of affairs itself. In other words, the object of a person's desire is a future state of affairs rather than the truth value of a proposition. When we consider cases involving future-directed desires, I submit that we need to focus on the cases in which we desire the obtaining of the future state of affairs itself. The problem with focusing on the truth value of propositions is that the desire is plausibly not actually *future-directed*. If Eve

56 This point is illustrated in the tenure example earlier. I am not tenured now even though (a) I have submitted my application, and (b) it is true that my application will be approved next month.

57 See also Bradley, "Well-Being at a Time," for a different argument based on truth makers that further supports this point (24).

desires that it is true *now* in 2025 that Eve gets pregnant in 2026, this desire is *present-directed*—the desire is directed towards what truth value the proposition *presently* has in 2025.⁵⁸ Insofar as what a person desires in a case involving future-directed desires is the future state of affairs itself, desire-satisfactionists should hold that the necessary condition of having a future-directed desire satisfied is (a) desiring *p* and (b) *p* obtains.

6. CAN PEOPLE BENEFIT IN CASES INVOLVING FUTURE-DIRECTED DESIRES?

My argument for no-future time-of-desire relies on the synchronic resonance constraint and on the all-conditions-met principle. However, one might object to my argument for no-future time-of-desire by pointing out that both the synchronic resonance constraint and the all-conditions-met principle address only the timing question. In cases involving future-directed desires, these principles jointly entail that a person benefits neither at t_d nor at t_o . However, these principles do not drive us to hold that people cannot benefit in such cases. We need to consider the possibility that in cases involving future-directed desires, a person benefits (1) at some other time, (2) across their whole life, (3) eternally, or (4) at no time.

I think that these options are all problematic. Consider option 1. The synchronic resonance constraint rules out all times Gary did not have the desire to live in Japan. The all-conditions-met principle rules out all times before he lived in Japan. So, there does not seem to be any other time that we can specify when Gary benefits except times when he had the desire and is living (or has lived) in Japan. Options 2 and 3 can also be ruled out. Since Gary does not have the desire across his whole life or eternally, the synchronic resonance constraint entails that he does not benefit across his whole life or eternally.

This leaves option 4. Since the two principles I use rule out the various times, one might find it appealing to hold that a person benefits atemporally in cases involving future-directed desires. This view—call it *future-atemporal time-of-desire*—would hence look like this:

- (a) In cases involving past-directed (and present-directed) desires, a person benefits at t_d ; and (b) in cases involving future-directed desires, a person benefits atemporally.

⁵⁸ Various philosophers hold that the truth value of propositions is not actually temporally indexed to the present. Some hold that propositions are true eternally, while others hold that propositions are true atemporally. However, even on such views, the desire that a proposition is true is not *future-directed*. They are *eternally* directed or *atemporally* directed.

However, I think future-atemporal time-of-desire faces at least three problems. First, such a view clashes with our intuitions regarding Living in Japan. In Living in Japan, it seems that Gary does not benefit at all, whether at any time or timelessly. Second, it would be odd if we hold that benefits are located temporally in cases involving past-directed and present-directed desires, but not in cases involving future-directed desires. What could possibly explain this asymmetry? Third, it is odd if a benefit is not located in time, even though both the future-directed desire and the desired object are located temporally. This calls out for an explanation.

One might object that no-future time-of-desire's different treatment of future-directed desires and past-directed desires also posits an odd asymmetry. In response, I agree that both no-future time-of-desire and future-atemporal time-of-desire look odd because they treat future-directed desires differently from past-directed and present-directed desires. However, the two earlier theoretical principles—synchronic resonance constraint and the all-conditions-met principle—are supposed to explain why an asymmetry is expected. So, the oddness of the asymmetry has an explanation.

In any case, I maintain that future-atemporal time-of-desire has *additional* oddness that these principles cannot help explain since its oddness is due not just to positing an asymmetry. The oddness is also about explaining why the benefit is atemporal even though the desire and the desired object are located temporally.

Given that options 1–4 fail, we can conclude that a person benefits neither at any point in time nor atemporally in cases involving future-directed desires. Hence, people do not benefit in such cases.

7. THE NORMATIVE SIGNIFICANCE OF POSTHUMOUS EVENTS

Next, one might object directly to no-future time-of-desire by arguing that the satisfaction of future-directed desires is relevant to well-being. One might argue for this by appealing to the normative significance of posthumous events. There is a widespread intuition that people should honor dead people's wills, people should not harvest a dead person's organs if they had objected to donating their organs, and so forth. Many of the desires people have while alive seem to have normative significance for others after their death.⁵⁹ In order to account for why such desires have normative significance for others, one might suggest that we should hold that posthumous events can be good or bad for a person.

59 See, for example, Bykvist, "The Moral Relevance of Past Preferences," 117–20.

So, future-directed desires people had while they were alive can contribute to their well-being if satisfied.

This is an interesting objection, and I cannot fully take it up in this article as the literature on the philosophy of death is complicated. Let me sketch one short reply. I think desire satisfactionists have good reason to think that people cannot be benefited by posthumous events. *When* could such events benefit a person? Trying to account for when posthumous events can benefit (or harm) a person runs into similar problems for those trying to account for the timing of death's badness.⁶⁰ If desire satisfactionists propose that posthumous events can benefit a person after they die, they face the same problem as those who hold that death is bad for a person after the person is dead. The problem is that neither the person nor their well-being level exists after they are dead.⁶¹

If desire satisfactionists propose that posthumous events can benefit a person before the person dies, they face the same problem as those who hold that death is bad for a person before the person is dead. The problem is that death does not seem bad for us while we are alive because it has not yet occurred. Similarly, it seems implausible to think that posthumous events can benefit us at a time when the event has not yet occurred. In addition, desire satisfactionists should hold that posthumous events cannot benefit a person prior to their occurrence as such benefits would violate the all-conditions-met principle, which I have defended.

One might instead propose that posthumous events can benefit a person atemporally instead. However, given my arguments in section 6, this seems like an implausible option for desire satisfactionists. If people can benefit in cases involving future-directed desires, desire satisfactionists should locate the benefit temporally. Given that desire satisfactionists should hold that posthumous events cannot benefit people before they die, after they die, or atemporally,

60 The problem of locating the time of benefit of posthumous events is even worse than locating the time of death's badness. In accounting for the badness of death, we might propose a deprivation account that appeals to comparative badness: our well-being level when dead is lower than our well-being level in the closest world in which we did not die. But when discussing desire satisfactionism, we are not talking about comparative benefits. To hold that our well-being is higher in a world where the posthumous event occurs than in a world where the event does not occur, we first need to show that the occurrence of the event benefits us in the former world.

61 Although most philosophers seem to hold this, some instead think that nonexistent people have a well-being level of zero permanently. See, for example, Bradley, *Well-Being and Death*, 98–111. This view faces various problems. See Ekendahl, "Responding to the Timing Argument," 765–66. More importantly, this view does not affect my argument—since the view holds that a dead person's well-being is always at zero after death, they are not (noncomparatively) benefitted or harmed after death.

desire satisfactionists have good reason to think that people cannot be benefited by posthumous events.

I suggest, then, that desire satisfactionists should look elsewhere to explain the normative significance of past desires that people had while they were alive. Indeed, Dorsey notes various problems with appealing to well-being to explain the normative significance of dead people's past desires and has proposed his own account that appeals to reasons of friendship instead of well-being.⁶² Fully figuring out why such desires are normatively significant if they cannot affect well-being is a task I leave for another paper.

8. SIMPLICITY AND RELATED WORRIES

Some critics say that my proposed view seems objectionably ad hoc, is non-parsimonious, is not simple, is less elegant, or seems to posit an odd and unexpected asymmetry.⁶³ How problematic are these sorts of worries for my theory? First, I reject the claim that the theory is objectionably ad hoc. Typically, a theory is objectionably ad hoc when it is unmotivated and/or tweaked just to handle special cases or counterexamples. However, I have motivated the theory by appealing to both intuitions and plausible principles. Furthermore, the theory is not proposed just to avoid some special cases or counterexamples.

Second, a common way of understanding parsimony and simplicity is in terms of Ockham's razor, according to which "entities are not to be multiplied beyond necessity."⁶⁴ Given Ockham's razor, we should prefer a more ontologically parsimonious theory, given that all other things are equal. Notice, however, that my theory does not violate Ockham's razor. No-future time-of-desire does not posit a greater number of types of entities (and also does not posit a greater number of entities). It is as ontologically parsimonious as its competitors.

Third, one might instead think that the problem with my proposal has to do with *syntactic* simplicity or elegance.⁶⁵ Since no-future time-of-desire posits an asymmetry between past-directed desires and future-directed desires, it is syntactically more complex than a number of its competitors, which treat both types of desires the same. I do not, however, think this is very problematic.

62 For example, Dorsey argues that many desires we have while alive are not relevant to our well-being but are instead relevant only to the well-being of others ("Friendship and the Wishes of the Dead," 139–32). Yet it seems that such desires have the same normative significance.

63 This objection has been raised by reviewers, at conference presentations, and in correspondence with other philosophers.

64 Baker, "Simplicity." See also Fitzpatrick, "Simplicity in the Philosophy of Science."

65 See Baker, "Simplicity"; and Fitzpatrick, "Simplicity in the Philosophy of Science."

Standardly, it is thought that we should prefer a simpler theory to more complex theories *all other things being equal*.⁶⁶ However, I have argued that things are not equal between the theories.⁶⁷

Finally, I have argued that there are two plausible principles that give us reason to expect an asymmetry. In past-directed cases, both conditions for benefitting are fulfilled only at t_d . Since the object resonates with a person at t_d , they can benefit then. In future-directed cases, both conditions for benefitting are fulfilled only at t_o . However, since the object alienates the person at that time, the person cannot benefit then. Given this, we should expect that the right theory is not simple and unified and that it would have this “odd” asymmetry. Hence, the lack of simplicity is not objectionable since it is explained by plausible principles.

9. BUT WE HAVE MANY FUTURE-DIRECTED DESIRES!

There is one final objection I want to address. In objecting to concurrentism, Bradley, Dorsey, and Purves each argue that because many of our desires are past-directed and future-directed, desire satisfactionists should think their satisfaction benefits us.⁶⁸ For example, Dorsey says, “Insofar as many of our evaluative attitudes are future-directed, one would expect a subjectivist theory to account for their prudential significance.”⁶⁹ Hence, concurrentism is problematic. This objection to concurrentism extends to no-future time-of-desire since no-future time-of-desire does not allow benefits in cases involving future-directed desires. Hence, no-future time-of-desire may seem similarly problematic.⁷⁰

66 Baker, “Simplicity”; and Fitzpatrick, “Simplicity in the Philosophy of Science.”

67 An objector might depart from standard formulations and hold that simplicity is still a virtue even when other things are not equal. Still, I think that it is unclear how much weight we should put on simplicity, and I am skeptical that simplicity should be given so much weight that it can trump the arguments and intuitions outlined in this article.

68 Bradley, *Well-Being and Death*, 23; Dorsey, “Desire-Satisfaction and Welfare as Temporal,” 157, and *A Theory of Prudence*, 194; and Purves, “Desire Satisfaction, Death, and Time,” 803.

69 Dorsey, *A Theory of Prudence*, 194.

70 One reviewer notes that a more charitable interpretation of their objection is not that the mere quantity of future-directed desires gives us reason to think they are relevant to well-being. Rather, the problem is that if we count the great number of future-directed desires as welfare irrelevant, it leads to counterintuitive results whereby a person’s life is not high in well-being even though their life is filled with a great amount of satisfaction of future-directed desires. In response, I do not think that it is intuitive to think that having a great amount of satisfaction of future-directed desires results in lots of benefits *given that the desired objects obtain at times when one has no desires for them*. We can imagine that a lot

I think this objection fails. First, the quantities of a certain kind of desire do not make a difference in whether satisfying those desires can contribute to well-being. Suppose everyone takes a drug that causes them not to care much about the present. Hence, people have few present-directed desires. Would the reduction in the quantity of present-directed desires make satisfying such desires irrelevant to well-being? It seems not. Intuitively, satisfying present-directed desires is relevant to well-being regardless of their quantity.

Second, we often have reasons to think that certain types of desires are irrelevant to well-being. Desire satisfactionists themselves have introduced various restrictions over which desires are relevant to well-being. For example, desire satisfactionists usually think that nonintrinsic desires are irrelevant to well-being because well-being is about what is *intrinsically* good for a person. So, even though we have a great number of nonintrinsic desires, there is a principled reason not to account for their welfare value. This is not the only example. Some desire satisfactionists argue that only idealized/informed desires or self-regarding desires are relevant to well-being. They argue that other types of desires are irrelevant to well-being (even though people have a great number of such desires). Similarly, given my arguments, we have a principled reason to think that future-directed desires are irrelevant to well-being despite a large number of such desires.

10. CONCLUSION

Let me summarize. First, I have argued that only no-future time-of-desire accommodates our intuitions in both *Singing* and *Living in Japan*. Second, I have defended the synchronic resonance constraint and the all-conditions-met principle and argued that no-future time-of-desire accommodates both principles. Although concurrentism can similarly accommodate both principles, I have argued that concurrentism has unintuitive results and objectionably restricts temporal distance while allowing spatial distance. Finally, I have argued that in cases involving future-directed desires, a person neither benefits at any point in time nor atemporally, and hence they do not benefit in such cases. Therefore, we have good reason to accept no-future time-of-desire over competing views.

One final thing to note. Although I have not discussed restricted forms of desire satisfactionism, I believe the cases and arguments apply to many of these other versions. The desires discussed in *Singing* and *Living in Japan* are

of cases similar to *Living in Japan* occur in a person's life. It does not seem like the person's life is high in well-being in virtue of these types of occurrences.

nonaltruistic, nonmoral, self-regarding, and plausibly directed towards fitting objects; and we can take Helen and Gary as being under idealized conditions (e.g., being fully rational adults at their different ages and aware of the relevant facts at each time). So, no-future time-of-desire can be applied to various restricted forms of desire satisfactionism. Similarly, the cases and arguments can also be framed in terms of other pro-attitudes (like valuing) instead of desires. Therefore, no-future time-of-desire can also be applied to other similar subjectivist theories of well-being.⁷¹

Rutgers–New Brunswick
fred-choo@outlook.com

REFERENCES

- Baber, H. E. “*Ex Ante* Desire and *Post Hoc* Satisfaction.” In *Time and Identity*, edited by Joseph Keim Campbell, Michael O’Rourke, and Harry S. Silverstein. Massachusetts Institute of Technology Press, 2010.
- Baker, Alan. “Simplicity.” In *Stanford Encyclopedia of Philosophy* (Summer 2022). <https://plato.stanford.edu/archives/sum2022/entries/simplicity/>.
- Bradley, Ben. *Well-Being and Death*. Clarendon Press, 2009.
- . “Well-Being at a Time.” *Philosophic Exchange* 45, no. 1 (2016): 1–12.
- Bramble, Ben. *The Passing of Temporal Well-Being*. Routledge, 2018.
- Bruckner, Donald W. “Present Desire Satisfaction and Past Well-Being.” *Australasian Journal of Philosophy* 91, no. 1 (2013): 15–29.
- Bykvist, Krister. “The Moral Relevance of Past Preferences.” In *Time and Ethics: Essays at the Intersection*, edited by Heather Dyke. Kluwer Academic Publishers, 2003.
- . “Sumner on Desires and Well-Being.” *Canadian Journal of Philosophy* 32, no. 4 (2002): 475–90.
- Campbell, Stephen M. “When the Shape of a Life Matters.” *Ethical Theory and Moral Practice* 18, no. 3 (2015): 565–75.
- Dorsey, Dale. “Desire-Satisfaction and Welfare as Temporal.” *Ethical Theory and*

71 I wish to thank Esther Goh, Zoë Johnson King, Connie Rosati, Alex Guerrero, Frances Kamm, Michael Otsuka, Carolina Sartorio, Aaron Thieme, Milan Ney, Kyle York, Felipe Pereira, Harisan Nasir, Itamar Saadon, Audrey Powers, Diego Arana, Rose Fonth, Éno Agolli, Kayla Jackson, Jordan Scott, Andrew Forechimes, Preston Greene, Andres Luco, and anonymous reviewers. I also like to thank participants at the sixth Beijing Analytic Philosophy Conference (2019), sixteenth Rocky Mountain Ethics Congress (ROME 2023), and APA Central Division Meeting 2024.

- Moral Practice* 16, no. 1 (2013): 151–71.
- . “Friendship and the Wishes of the Dead.” *Legal Theory* 28, no. 2 (2022): 124–45.
- . *A Theory of Prudence*. Oxford University Press, 2021.
- . “Why Should Welfare ‘Fit’?” *Philosophical Quarterly* 67, no. 269 (2017): 685–708.
- Ekendahl, Karl. “Responding to the Timing Argument.” *Pacific Philosophical Quarterly* 103, no. 4 (2022): 753–71.
- Fischer, John Martin. *Our Stories: Essays on Life, Death, and Free Will*. Oxford University Press, 2009.
- Fitzpatrick, Simon. “Simplicity in the Philosophy of Science.” In *Internet Encyclopedia of Philosophy*. <https://iep.utm.edu/simplici/>.
- Forrester, Paul. “Concurrent Awareness Desire Satisfactionism.” *Utilitas* 35, no. 3 (2023): 198–217.
- Heathwood, Chris. “Desire Satisfactionism and Hedonism.” *Philosophical Studies* 128, no. 3 (2006): 539–63.
- . “The Problem of Defective Desires.” *Australasian Journal of Philosophy* 83, no. 4 (2005): 487–504.
- Johansson, Jens. “The Time of Death’s Badness.” *Journal of Medicine and Philosophy* 37, no. 5 (2012): 464–79.
- King, Owen C. “The Good of Today Depends Not on the Good of Tomorrow: A Constraint on Theories of Well-Being.” *Philosophical Studies* 177, no. 8 (2020): 2365–80.
- . “Pulling Apart Well-Being at a Time and the Goodness of a Life.” *Ergo* 5 (2018): 349–70.
- Lin, Eden. “Asymmetrism About Desire Satisfactionism and Time.” In *Oxford Studies in Normative Ethics*, vol. 7, edited by Mark Timmons. Oxford University Press, 2017.
- Portmore, Douglas W. “Desire Fulfillment and Posthumous Harm.” *American Philosophical Quarterly* 44, no. 1 (2007): 27–38.
- Purves, Duncan. “Desire Satisfaction, Death, and Time.” *Canadian Journal of Philosophy* 47, no. 6 (2017): 799–819.
- Railton, Peter. “Facts and Values.” In *Facts and Values and Norms: Essays Toward a Morality of Consequence*. Cambridge University Press, 2003.
- Rosati, Connie S. “Internalism and the Good for a Person.” *Ethics* 106, no. 2 (1996): 297–326.
- Sarch, Alexander. “Desire Satisfactionism and Time.” *Utilitas* 25, no. 2 (2013): 221–45.
- Silverstein, Harry S. “The Time of the Evil of Death.” In *Time and Identity*, edited by Joseph Keim Campbell, Michael O’Rourke, and Harry S. Silver-

stein. Massachusetts Institute of Technology Press, 2010.
Sumner, L. W. *Welfare, Happiness, and Ethics*. Clarendon Press, 1996.

DELEGITIMIZING TRANSPHOBIC VIEWS IN ACADEMIA

Logan Mitchell

ACADEMIC FREEDOM and social equality are cornerstones of university life, cherished by many. Roughly, academic freedom involves measures to protect members of an academic community from undue social or political influence in the context of their research and teaching. An academic community is harmed by both undue restrictions of such freedom and the perpetuation of unjust social hierarchies. Unsurprisingly, then, things get tense when these two values come into conflict within academic institutions such as research universities, community colleges, and university presses.

Such conflicts appear to force us to pick a side, and many oblige. For some, academic freedom takes precedence over social equality. Perhaps this value schema best protects the marginalized in the long run, or perhaps academic freedom is just more important. Others counter that social equality is a prerequisite for genuine academic freedom, and without the former, the latter is a chimera. Of course, many simply feel stuck in a dilemma with no good option: become an enemy of academic freedom or an inhibitor of equality. What should we do?

Views that deny transgender people's claims to equal standing or otherwise warrant the disrespect of transgender people present us with important instantiations of the conflict between academic freedom and social equality with notable political implications, warranting further exploration. Many people, for example, defend views that imply trans people deserve less social or political power or respect than cisgender people.¹ I refer to such views as *transphobic*. Some transphobic views are clearly worthy of sanction, like the idea that being

1 In this article, I use 'trans' to refer to those whose gender modality is transgender. As Florence Ashley, Shari Brightly-Brown, and G. Nic Rider argue, someone's gender modality refers to how that person's gender identity relates to their gender assigned at birth ("Beyond the Trans/Cis Binary"). Transgender people are those whose gender identities do not correspond to their genders assigned at birth. Of course, the very concept of gender identity is subject to critical scrutiny. See Ashley, "What Is It Like to Have a Gender Identity?"; Andler, "Gender Identity and Exclusion"; Clare, *Nonbinary*; and Hernandez and Bell, "Much Ado About Nothing." When speaking of gender identity in this article, I follow E. M. Hernandez and Rowan Bell's deflationary account of gender identity, according

trans warrants assault. Others, however, are perceived to be more nuanced (at least to some parties) and less worthy of sanction. Take the stance of many “gender-critical feminists” who claim that while trans people deserve basic respect and care, their claims to gender self-determination should be rejected, along with the entitlements such self-determination would carry with it, such as access to certain gendered spaces.²

These supposedly nuanced transphobic views pose a unique problem for academics who cherish both academic freedom and social equality. It is generally uncontested whether, for example, explicitly white supremacist views ought to be published in prestigious academic journals: there is enough consensus, at least within academia, that such views are not viable candidates for belief (not to mention that they are harmful).³ However, trans-exclusionary views have recently been published in well-respected venues, and when trans-exclusionary submissions have been rejected, claims of unjust censorship are easy to find.⁴

Considerations of academic freedom seem to compel providing platforms of some sort to those advocating for supposedly nuanced transphobic views, even if those views are implausible and harmful.⁵ Considerations of social equality, on the other hand, appear to speak in favor of institutionally affirming the validity of trans people’s genders, which in turn calls for the explicit rejection of views that perpetuate the subordinate social status of trans people.⁶ Given this push and pull, it should not come as a shock that there exists significant disagreement about how we ought to respond to supposedly nuanced transphobic views from inside (and outside) academia.

Despite the appearance of a dilemma, I am optimistic that a path exists that, if successful, allows academic institutions to sufficiently honor both social equality and academic freedom. This path relies on distinguishing between a

to which one’s gender identity is “just the gender one identifies as” (“Much Ado About Nothing,” 3).

2 See, e.g., Lawford-Smith, *Gender-Critical Feminism*; and Stock, *Material Girls*.

3 By ‘explicitly white supremacist view,’ I mean to refer roughly to views that imply nonwhite people deserve less social or political power or respect than white people. Even on this rough view, however, there have been some plausibly white supremacist views published in academic journals, such as Bruce Gilley’s highly controversial paper “The Case for Colonialism,” which was ultimately retracted from its original journal. However, given the rarity of such examples and the outcry they reliably produce, I take it that such views are still generally taken to be institutionally illegitimate by most academics.

4 See Lawford-Smith, *Gender-Critical Feminism*; and Byrne, “Philosophy’s No-Go Zone.”

5 See “Open Letter of Solidarity with the University of Sussex from UK Philosophers,” October 11, 2021, <https://openlettertosussexfromukphilosophers.wordpress.com/>.

6 See “Open Letter Concerning Transphobia in Philosophy,” January 2020, <https://sites.google.com/view/trans-phil-letter/>.

view being *censored* and its being *institutionally delegitimized*. While a detailed exploration is forthcoming below, the rough idea is that a view is institutionally illegitimate when institutional actors do not consider the view to be worth taking seriously as a viable candidate for belief. To delegitimize a view is to take steps to ensure the view is not taken seriously as a viable candidate for belief by the institution.

In this article, then, I argue that academic institutions have a *pro tanto* obligation to *delegitimize* all transphobic views, and in many contexts, the obligation is undefeated. By this, I mean that academic institutions generally should not take such views seriously as viable candidates for belief, though in some cases, this obligation may be outweighed by other considerations. Three premises together justify this conclusion. First, if academic institutions do not delegitimize transphobic views, then they structurally perpetuate the subordination of trans people. (Institutions perpetuate subordination through their own practices and policies.) Second, institutions have a *pro tanto* obligation to avoid structurally perpetuating subordination, and this obligation can be defeated only when such avoidance is excessively burdensome. Third, academic institutions can delegitimize transphobic views in a manner that is not excessively burdensome, at least in many contexts. When it comes to academic institutions, plausibly the two most relevant burdens to avoid are unduly restricting academic freedom and compromising the integrity of the institution in question, so when focusing on the possible burdens of delegitimization, I focus on these two issues.⁷ While a similar argument could be made for nonacademic institutions like the news media, I choose a somewhat narrower scope so that I can focus on the unique nature of academic freedom, which is importantly distinct from free speech more broadly construed.⁸

I take the second premise of my argument to be relatively uncontroversial and so will not spend time defending it. When an institution is perpetuating

7 There are other burdens that can affect an institution's obligation to delegitimize. Administrative burdens or donor interests, for example, can be such that delegitimization constitutes an excessive burden in some cases. However, I suspect that considerations of academic freedom and institutional integrity pose the strongest challenges to delegitimization on philosophical grounds. As such, in this article, I choose to focus on only these two burdens.

8 For more on the difference between academic freedom and free speech, see Simpson and Srinivasan, "No Platforming"; Moody-Adams, "What's So Special About Academic Freedom?"; Van Alstyne, "The Specific Theory of Academic Freedom and the General Issue of Civil Liberty"; Simpson, "The Relation Between Academic Freedom and Free Speech"; Post, *Democracy, Expertise, and Academic Freedom*; and Fish, *Versions of Academic Freedom*. A detailed exploration of the differences between academic freedom and free speech can also be found in section 3.2 below.

the subordination of a group through its own doing, it has a *pro tanto* obligation to stop doing so. While this obligation may be defeated when stopping is too costly or strays too far from the institution's aims, it is still weighty—an institution must have a strong reason to continue operating in an unjust manner. I am not claiming that this obligation exhausts institutions' duties to redress subordination. It may be the case that institutions ought to stop perpetuating subordination *even if* it is very burdensome or that institutions have a broader obligation to redress more distant subordination. All that is required to motivate my conclusion, however, is the weaker claim that institutions ought to stop engaging in behaviors that structurally perpetuate a group's subordinate social status when stopping would not be too burdensome.

While it may be easy to accept that institutions ought to avoid structurally perpetuating subordination, determining the best way to ameliorate such subordination in the case of trans people is more controversial. In this article, I assume that trans people have the right to *gender self-determination* and are entitled to the goods and services that arise out of this right. In discussing the right to gender self-determination, I follow Florence Ashley, who argues that the right to gender self-determination “means that individuals have a right to define, express, and embody their gender identity as they see fit . . . and can be derived from and is supported by many long-recognized rights, including the right to free speech, equality, privacy, identity, and dignity, and to live and act with integrity.”⁹ Some of the entitlements that plausibly arise out of this right include having one's government identification match one's stated gender, access to gender-specific spaces that align with one's gender, affordable access to at least some forms of gender-affirming health care, and the ability to pursue legal recourse against gender-based discrimination in employment and housing.

Given my assumption that trans people have the right to gender self-determination, I further assume that to refuse trans people this right is in fact to deny trans people social and political power to which they are entitled, which in turn perpetuates trans people's undeserved subordinate social status. I acknowledge that such assumptions are contested, and what follows is unlikely to convince those who are not trans affirming.¹⁰ However, this article does not merely

9 Ashley, “Gender Self-Determination as a Medical Right,” 833.

10 As Ashley notes, a gender-affirmative approach “affirms the individual's gender socially and medically” (“Homophobia, Conversion Therapy, and Care Models for Trans Youth,” 361). Similarly, Hernandez argues that gender affirmation involves *perceiving* an individual on their own terms, which involves giving uptake to what Robin Dembroff and Cat Saint-Croix call someone's *agential identity*—“the part(s) of one's self-identity that one presents to the world that connects with their social position.” See Hernandez, “Gender-Affirmation and Loving Attention,” 8; and Dembroff and Saint-Croix, “Yep, I'm Gay.” Thus, in

preach to the choir. For many people who are trans affirming may nevertheless doubt that we should institutionally delegitimize transphobic views. Even if one grants the second premise of my argument, for example, the first and third premises may still raise eyebrows, particularly among those with strong Millian inclinations. There are also sure to be questions about what it means, exactly, to delegitimize a view.

So to begin, in section 1, I present an account of institutional legitimacy and briefly explore how one can practically delegitimize a view. After laying out my proposal, I begin to defend my first premise in section 2, drawing on the work of Sophia Moreau to argue that the institutional legitimacy of transphobic views structurally perpetuates the subordination of trans people.¹¹ This section comes with an account of what makes a view transphobic, to make clear the set of views I consider viable candidates for delegitimization. Then I defend the third premise in section 3, arguing that delegitimization, at least in many contexts, is not nearly as radical as some might think, nor does it run afoul of a robust notion of academic freedom. Along the way (section 3.2), I finish defending the first premise by addressing the objection that there might be viable alternatives to delegitimization.

While this article focuses on transphobic views, the argument presented has broad implications about how we ought to balance academic freedom and social equality, particularly regarding views that exacerbate a social group's subordination if given uptake. Delegitimization thus presents itself as a solution to similar conflict cases. At the same time, delegitimization as argued for here can avoid slippery-slope objections that warn of the potential demise of open discourse. As I show, my proposal does not entail banning discussion of abhorrent views in most institutional contexts. Additionally, what warrants delegitimization on my account is a clear link between legitimacy and subordination—not feelings of offense, moral repugnance, or political unpopularity. So my argument does not speak in favor of broad censorship of politically unpopular views. Delegitimization, then, is a flexible and reasonable solution that poses no threat to the flourishing of truth-seeking institutions.

1. INSTITUTIONAL LEGITIMACY

I claim that we ought to institutionally delegitimize transphobic views. In this section, I clarify this proposal by explaining what makes a view institutionally

speaking of a person's or an institution's being trans affirming, I refer to a commitment made to affirm trans people's stated genders, which can manifest on a variety of levels from institutional policies to interpersonal practices.

11 Moreau, "Discrimination and Subordination."

legitimate (section 1.1) and discussing what delegitimization might look like in practice (section 1.2). As I show, delegitimization can take many forms, some more radical than others.

1.1. *Institutional Legitimacy*

When it comes to views, we can speak of their possessing different types of legitimacy. By *view*, I mean truth-evaluable propositions, whether a single idea (e.g., the earth is flat) or a particular collection of ideas (e.g., an argument with the conclusion that the earth is flat). Normatively, we might talk about what I will call *epistemic* legitimacy. An epistemically legitimate view is, roughly, one that is worth taking seriously as a viable candidate for belief, relative to a community at a given point in time. Epistemic legitimacy is thus a purely *normative* category—such views warrant consideration based on their merits and the available evidence, regardless of whether they are, in practice, given such consideration. Contradictory views can be epistemically legitimate, such as those for which evidence supports different conclusions. As such, not all epistemically legitimate views are true. Epistemic legitimacy can also change over time as new evidence becomes available. The view that the earth is flat may have been worth taking seriously at one point in time, but subsequent evidence has since undermined its legitimacy. Nowadays, flat earth views just are not worth taking seriously as viable candidates for belief.

By contrast, we can speak of various kinds of *social* legitimacy. A view is socially legitimate if there are social practices in place that uphold the notion that the view is worth taking seriously as a viable candidate for belief. While epistemic legitimacy is a normative category, social legitimacy and its various subsets are *descriptive* categories grounded in the normative attitudes of a group's members. Consequently, socially legitimate views can be and often are epistemically illegitimate. Patently false antivaccine conspiracy theories are taken seriously among many social groups, for example, and are socially legitimate for them. There are various contexts in which a view can be socially legitimate. As such, views that are socially legitimate in one context may be socially illegitimate in another. While antivaccine conspiracy theories are socially legitimate in certain circles, such views are not socially legitimate amongst epidemiologists.

Institutional legitimacy is a subset of social legitimacy that can describe either the normative attitudes that members of an institution hold or certain institutional structures that signal that a view is a viable candidate for belief. A view is institutionally legitimate, then, when either: (i) the view is considered a viable candidate for belief by enough members of a particular institution, and there are various practices or policies in place to perpetuate this notion; or (ii)

the institution is structured in such a way that signals that the view is a viable candidate for belief (regardless of how many members of the institution actually consider it to be legitimate). In some cases, institutional decision-makers (e.g., deans or university presidents) may have enough power to make a view institutionally legitimate *even if* very few institutional members consider the view to be a viable candidate for belief. Thus, institutional legitimacy can arise through a variety of means, some more democratic than others. Institutional legitimacy also comes in degrees, with some views being taken more seriously than others.

It is worth emphasizing at the outset that all institutions consider some views to be institutionally illegitimate, and most people do not find this fact alone objectionable. For any institution, no matter its goals, some views just are not considered worthy of serious epistemic consideration by the institution's own lights.¹² Institutional members have limited resources, and some inquiries have simply been closed. Of course, whether a particular view *ought* to be ascribed legitimacy is often hotly contested. Furthermore, institutions can deny legitimacy to a view when no genuine justification exists, such as attempting to suppress views simply because they are politically unpopular. As such, there are also plenty of cases of delegitimization that are rightfully criticized as pernicious.

Institutional legitimacy does not neatly track social legitimacy broadly construed (i.e., public opinion), nor should it. Most truth-seeking institutions, for example, do not ascribe legitimacy to the view that Donald Trump won the 2020 election, even though as of September 2022 over 60 percent of Republicans and 25 percent of Americans still believed this claim.¹³ When it comes to vaccine efficacy, election security, and climate change, most American truth-seeking institutions do not include antivaxxers, Trumpian conspiracy theorists, or climate deniers on their panels of experts, despite the large number of Americans who hold such views. Such views are not considered viable candidates for belief, and advocates of such views are not considered competent experts.

Similarly, institutional legitimacy need not always aim to track epistemic legitimacy. An institution might have a legitimate goal that requires that certain views be rejected *regardless* of their viability. As I argue in section 3, for example, goals like racial integration or combatting misogyny might demand that some views be institutionally delegitimized, even if some institutional members consider such views to be epistemically legitimate or if the institution

12 Simpson and Srinivasan, "No Platforming."

13 Bump, "Nearly 700 Days Later, Most Republicans Still Believe Trump's Big Lie"; and Kahn, "53% of Republicans View Trump as True US President."

might normally remain neutral about such matters. In such cases, a view's epistemic status may be deemed irrelevant in determining how to treat it. However, many institutions (particularly those in the business of producing knowledge) believe it justified to delegitimize only views that are epistemically illegitimate and thus aim for maximal alignment between institutional legitimacy and epistemic legitimacy.

To understand institutional legitimacy in action, consider certain views about interracial marriage. In the United States, citizens are legally free to argue that interracial marriage is wrong because white people are superior to other races. In fact, as of 2021, 7 percent of white Americans still disapproved of interracial marriage.¹⁴ However, such views are not legitimate in most institutional settings, whether they be academic, corporate, or philanthropic. It is difficult to fathom a prestigious philanthropic foundation providing grants to organizations that publicly advocate for stripping legal rights from those involved in interracial marriages. Recent examples of such white supremacist arguments cannot be found in credible academic journals.

In other words, white supremacist opposition to interracial marriage is *institutionally illegitimate* in most American institutions, even though at one point, this was not the case. Over time, institutional actors within academia, the news media, and other organizations have come to consider such opposition to be unworthy of sincere epistemic engagement. Furthermore, most institutional members are on board with such views' illegitimacy—it is not common to hear people criticize such illegitimacy as a stifling of academic freedom, which I discuss in more detail below in section 3.2.

1.2. *On the Treatment of Institutionally Illegitimate Views*

So, institutionally illegitimate views are not considered to be worth taking seriously as viable candidates for belief. Let us now discuss, practically, how institutionally illegitimate views might be treated. Since there are competing, reasonable approaches regarding exactly how to best treat such views, I provide a general sketch without privileging one approach over another.

Let me begin by emphasizing that institutionally illegitimate views are not usually flat-out banned or even discouraged from being discussed. In fact, there are often valuable reasons for discussing such views. On the night of the 2020 presidential election, many news organizations aired a speech in which Donald Trump claimed he had won the election; however, most of the organizations were quick to clarify that such claims were false, with discussions focusing on

14 McCarthy, "US Approval of Interracial Marriage at New High of 94%."

the threat that such lies pose to American democracy.¹⁵ Trump's claims of winning were denied institutional legitimacy, even though they were discussed extensively.¹⁶

This illustrates that one might take a view seriously as a vehicle of harm, as a useful pedagogical tool, or as a necessary inconvenience—all without taking the view seriously *as a viable candidate for belief*. Similarly, *debunking* a particular view for epistemic, political, or moral purposes can occur without ascribing the view credibility. It can be beneficial to engage with white supremacist rhetoric or flat-earth theories to point out their flaws and show precisely why they are illegitimate. Thoroughly analyzing certain views may even be helpful for members of marginalized communities to counteract particular views (even if it is unjust that they must do such counteracting in the first place). So to institutionally delegitimize a view is compatible with the view's being discussed often, even within various institutional contexts.

However, because institutionally illegitimate views are not ascribed *credibility* by institutional members, there are reasons to treat them differently from institutionally legitimate views in at least some contexts. In academia, some means of expressing a view ascribe credibility to it while others do not. How institutionally illegitimate views are treated in a university thus depends on the context of expression. To illustrate this, I briefly touch on three prominent domains of action in academia that plausibly call for asymmetry in the treatment of legitimate and illegitimate views—publishing, hiring, and (at least some forms of) platforming.

When it comes to publishing, institutionally illegitimate views are not considered suitable for being advocated for or endorsed in academic books, peer-reviewed articles, or conference presentations (though such views might be discussed at length).¹⁷ Accepting an article for publication, at least in many

15 For example, see Goodykoontz, “Why Trump’s Fake Victory Speech Was the Worst and Most Dangerous TV of Election Night 2020.”

16 Of course, advocates of such views do not merely wish to have their views discussed—they want their views to be legitimized. As I argue in section 3.2, however, academic freedom does not require that we provide such legitimacy to all those who ask for it.

17 One might argue that any form of engagement with a view, even when arguing against it, entails that one is treating the view as a viable candidate for belief (and thus legitimizing it). I disagree. That a view is being discussed simply entails that the discussant sees value in directly addressing the view in question. It might be valuable for a public health official to directly address the myth that vaccines cause autism; however, by addressing such myths, the public health official does not thereby imply that such views are more credible than other conspiracy theories. Similarly, authors like Robin Dembroff see value in directly addressing trans-exclusionary views (“Escaping the Natural Attitude About Gender”). However, to directly engage with a view in order to debunk it does not imply

contexts, carries with it an undeniable ascription of credibility—for an article to be published in a prestigious academic venue says something about the legitimacy of the views being endorsed in the article at the time of publication. Thus, if a journal receives a submission advocating for an institutionally illegitimate view (e.g., a white supremacist argument against interracial marriage), then by the institution's own lights, the view's illegitimacy is a reason for rejecting it.

Institutionally illegitimate views are also considered unfit to serve as the research aim of a professor, given a university's interest in ensuring standards of quality control. In the context of a hiring process for the position of a tenure-track professor, if a candidate has research that advocates for institutionally illegitimate views, then such advocacy counts against the candidate, if not excluding them from consideration altogether. Hiring someone to perform research or teach classes in a particular domain of expertise ascribes credibility to the candidate's views in that domain, and institutions might want to avoid ascribing credibility to the views being advocated by certain candidates.

Providing an institutional platform for speakers can also call for asymmetrical treatment between institutionally legitimate and illegitimate views. Consider, for example, the practice of *no-platforming*, which Michael Veber refers to as “preventing or prohibiting someone from contributing to public discussion because that person advances ... objectionable views,” which, in academic settings, “amounts to preventing someone from publicly speaking either by disrupting the event or refusing to allow it to be scheduled in the first place.”¹⁸ There are reasonable arguments for and against the practice of no-platforming, given the complex relationship that such practices have with regard to academic freedom, free speech, and other institutional norms.¹⁹ Furthermore, different platforms carry with them varying amounts of credibility—invitations sent by student groups might not carry the same weight as those sent by faculty or administration. Thus, some institutions might deem it acceptable to allow institutionally illegitimate views to be platformed in a limited set of cases. While I do not advocate for a particular position on the relationship between delegitimization and no-platforming, I mention the practice here to sketch a picture of how institutionally illegitimate views might be treated differently from those that are institutionally legitimate in common university contexts.

Also relevant for our purposes is the question of what to do when a view is currently institutionally legitimate but should not be, as I take to be the case

an endorsement of the view as epistemically legitimate, the way that advocating for the view or, for example, including defenses of the view in an edited volume does.

18 Veber, “The Epistemology of No Platforming,” 1.

19 See Simpson and Srinivasan, “No Platforming”; Veber, “The Epistemology of No Platforming”; and Levy, “No-Platforming and Higher-Order Evidence, or Anti-Anti-No-Platforming.”

with transphobic views (at least for many academic institutions). When I speak of an institution being obligated to delegitimize a particular view, I am talking about a *pro tanto* obligation held by an institution to alter its policies and practices to clearly reflect the notion that the view in question is not a viable candidate for belief. One might wonder whether there is a difference between being obligated to delegitimize a view and being obligated to simply *not* legitimize a view. Such a difference can be relevant when a view is already institutionally illegitimate—in such cases, the relevant institution need only *maintain* its practices and policies to ensure the view does not become legitimized. Importantly, however, for views that are already institutionally legitimate, the only way for an institution to discharge an obligation to not legitimize the view is in fact to delegitimize it—to take steps to ensure the view is clearly not taken seriously as a viable candidate for belief by the institution. If a view is legitimate, institutional neutrality merely perpetuates this legitimacy. Thus, in such cases, the institutional obligation to not legitimize a view has the same scope as the obligation to delegitimize it.

With that in mind, the distinction between delegitimizing and not legitimizing is relevant when considering that an institution's obligation to delegitimize a view creates a set of corollary obligations for institutional actors that can vary in scope. In some cases, certain actors, particularly those with notable decision-making authority, are obligated to take steps to institutionalize the illegitimacy of the view in question. This corollary obligation to *actively* delegitimize might be contrasted with a narrower obligation held by other institutional actors to merely not legitimize a view, which requires them only to refrain from further legitimizing the view, without taking steps to institutionalize this perspective. There may be cases where the narrower obligation is all that can be required of an institutional actor, but I assume that in many cases, the more active obligation holds, especially for those with significant institutional authority.²⁰

2. TRANSPHOBIC VIEWS, SUBORDINATION, AND STRUCTURAL ACCOMMODATION

Having explained what institutional legitimacy is, in this section I begin to defend the first premise of my argument: if an academic institution does not delegitimize transphobic views, then it structurally perpetuates the

20 Lastly, delegitimizing *views* is quite different from delegitimizing *people*. One can delegitimize a view without believing that advocates for illegitimate views ought to be kept out of institutions in any and all capacities, treated disrespectfully, or denied any credibility whatsoever.

subordination of trans people. I first provide an account of what makes a view transphobic (section 2.1), explicating the set of views that I take to be viable candidates for delegitimization. I then introduce Moreau's account of social subordination (section 2.2) to argue that the institutional legitimacy of transphobic views works in tandem with negative stereotypes about trans people to perpetuate their lack of social and political power relative to cisgender people who are similarly socially situated (section 2.3).

The primary aim of this section is to present the centerpiece of my defense of the first premise: the claim that the institutional legitimacy of transphobic views structurally perpetuates the subordination of trans people. Of course, fully defending the first premise requires more than this claim, as the first premise implies that academic institutions continue to perpetuate such subordination unless they *delegitimize* transphobic views. Some may object to this implication, arguing that there are other effective strategies to stop perpetuating such subordination that do not involve delegitimization. To sufficiently address this objection, I save discussion of it for the next section (section 3.2), at which point I take the first premise to have been fully defended.

2.1. *What Is a Transphobic View?*

While many commentators focus on transphobic actions, people, or attitudes, here I focus exclusively on views (i.e., individual propositions and collections of propositions). I consider a view transphobic when it expresses or implies that trans people deserve either less social or political power (when compared to cisgender people who are similarly socially situated) or disrespect. More specifically, I consider a view transphobic when it expresses or implies that agent *A*'s being transgender (i) provides justificatory support for *A*'s receiving less social or political power than were *A* cisgender or (ii) warrants treating *A* with greater censure or lesser consideration (i.e., disrespect) than were *A* cisgender.²¹ I call these the *disempowerment* and the *disrespect* conditions, respectively. Either condition alone is sufficient for a view's being transphobic, and many views satisfy both.

Throughout this discussion, I employ a distinction between transgender and cisgender people. It is important to note, however, that the distinction itself is subject to critical reflection, and in using it, I do not thereby endorse a simplistic binary between cisgender and transgender people.²² While cisgen-

21 This account is loosely inspired by Moreau's discussion of discriminatory policy in "Discrimination and Subordination," 121–22.

22 For more engagement with the term 'cis' and the cis/trans distinction, see Aultman, "Cisgender"; Serano, *Whipping Girl*; Darwin, "Challenging the Cisgender/Transgender Binary"; and Zurn, "Crippling Cis."

der and transgender are among two of the most common gender *modalities*, as Ashely argues, these two categories do not exhaust people's experience of gender.²³ Furthermore, in making comparisons between transgender and cisgender people, I invoke the notion of someone's being *similarly socially situated*, given that many trans people can still occupy positions of privilege along a variety of axes, including (but not limited to) class, race, and disability.

Let us now say more about these conditions and some of the views that plausibly satisfy them. The *disempowerment* condition states that a view is transphobic when it implies that being trans (or not being cisgender) warrants others limiting one's social or political power. The expression of such a view need not actually reduce trans people's power in order to meet the disempowerment condition. Rather, all that is needed is that the view lends justificatory support to such disempowerment. While questions about the nature of power are tricky, and I make no attempt to answer them here, I follow Moreau in holding that diminished power in the relevant sense can manifest as a reduction of freedom or *de facto* authority, of which the latter involves "a broader set of powers, including the power to be listened to, to be taken seriously when one brings a complaint against another."²⁴

Views that imply that trans women's political claims to womanhood should be rejected but cisgender women's should not be rejected meet the disempowerment condition. A trans woman presents a political claim to womanhood—she wants to be considered a woman by her society and government—and to deny this claim is to grant her less social and political power than were she a cisgender woman of a similar social situation. The same can be said of trans men and nonbinary people. Why does this denial amount to a reduction in social and political power? Cisgender people are free to use the restroom congruent with their gender identities, have legal recognition of their identified genders, and can access a variety of other legal entitlements. In stating that trans people's claims to such equality are unreasonable, one claims that trans people should *not* be free to move through the social and political world *as* their identified genders, while cisgender people are free to do so. In speaking of a reduction of power, then, I am speaking of power *relative* to a cisgender person who is similarly socially situated making the same claim. The view that trans people's testimony is less credible (perhaps in virtue of mental illness or extreme confusion) than that of cisgender people also meets the disempowerment condition.²⁵ On

23 Ashley, "Trans' Is My Gender Modality."

24 Moreau, "Discrimination and Subordination," 128.

25 Such treatment is often referred to as *testimonial injustice*. See Fricker, *Epistemic Injustice*.

such a view, an agent's being trans justifies limiting their *de facto* authority, since one would take trans people's claims and complaints less seriously.

Importantly, many views that are supposedly more nuanced also meet the disempowerment condition. Take, for example, the view that all trans people are, at best, *reasonably mistaken* about their gender identity, while cisgender people are not.²⁶ Perhaps trans people are generally credible knowers, but they have all simply made a mistake about their gender—only cisgender people correctly know their own genders. However, even reasonable mistakes do not justify granting someone legal entitlements, such as the right to use a single-gender restroom. So a trans man's being reasonably mistaken about being a man implies that his political claim to using the restroom is on less stable ground than were he cisgender. If I make a political claim (e.g., to use the men's restroom), and this claim relies upon a false premise (e.g., that I am a man), then my claim is weaker than a claim resting upon a true premise, if my claim has any weight at all. Similarly, I suspect that the view that a person is a woman if and only if they are an adult human female, where female is considered an immutable biological category, also meets the disempowerment condition.²⁷ This view often makes no explicit mention of transgender people but implies *at best* that all trans women are reasonably mistaken about their gender and implies that cisgender women should be entitled to more social and political power than transgender women.

Many gender-critical feminists object to the above reasoning. Kathleen Stock, for example, publicly expresses support of trans women being legally recognized as women in the UK, at least under certain conditions (e.g., having gender-affirming medical treatments and a formal diagnosis of gender dysphoria), while denying that trans women are *actually* women.²⁸ For Stock, granting a trans woman legal womanhood is to engage with a fiction, though such engagement can be socially useful. Thus, Stock and others are likely to maintain that holding a trans-exclusionary metaphysical view about womanhood does not necessarily meet the disempowerment condition.

Furthermore, Stock claims to object more specifically to "the institutionalization of the idea that *gender identity* is all that matters—that how you identify *automatically* confers all the entitlements of that sex."²⁹ When taken in isolation, this latter view is not necessarily transphobic since it is coherent to believe

26 I take this view to be among the most charitable positions that a trans-exclusionary person could take.

27 This view is analyzed thoroughly by Dembroff, "Escaping the Natural Attitude About Gender."

28 See Stock, "Response to Christa Peterson's Blog" and *Material Girls*, 178.

29 Moorhead, "Kathleen Stock" (emphasis added).

that trans people's claims to gender self-determination ought to be respected while also holding that, logistically, *some* requirements for legal recognition other than stated self-identification are warranted. However, Stock is ultimately committed to holding that all cisgender women are entitled to more political freedoms than trans women (particularly trans women who do not pass as cisgender or who have not had gender-affirming surgery). For example, Stock claims that when she refers to trans people with their pronouns of choice, she is simply being polite—for Stock, trans people have no *entitlement* to such treatment, let alone to access certain gendered spaces or other gender-based privileges.³⁰ Presumably, Stock believes that she herself is entitled to being referred to and viewed as a woman, both socially and legally. After all, she wants access to women-only spaces and other gender-based privileges.

To deny a trans person the same entitlements as a cisgender person in virtue of their gender identity is to give the trans person less social and political power. It is helpful here to recall the evolution of LGBTQ+ marriage rights in the United States. Having access to civil unions while being denied the full entitlements of marriage was inimical to the goal of affirming the social equality of heterosexual and nonheterosexual couples.³¹ Civil unions literally came with fewer political entitlements than marriage, and the mere distinction between the two affirmed that nonheterosexual couples were importantly (and to many, objectionably) distinct from heterosexual couples. While there are important disanalogies between LGBTQ+ marriage and trans rights, both cases involve an unequal distribution of political entitlements, with a subordinate group receiving less than a dominant group, where such a distribution is rightfully thought to run counter to social equality. So giving trans people some political entitlements while giving cisgender people more of those entitlements inhibits the ability of trans and cisgender people to relate to each other as equals, which many have argued is at the heart of social equality.³² Thus, merely to claim that cisgender women are actually women while trans women are not is plausibly transphobic on my account, since such claims justify the disempowerment of trans people.

The disrespect condition holds that being transgender warrants greater censure or lesser consideration than being cisgender. For example, the view that transgender people are sexually deviant and deceitful meets this condition since such traits mark trans people as uniquely worthy of censure.³³ Views that meet the disrespect condition almost always meet the disempowerment

30 Stock, *Material Girls*, 178. See Briggie, "Which Reality?"

31 Card, "Gay Divorce"; and Calhoun, *Feminism, the Family, and the Politics of the Closet*.

32 See, e.g., Anderson, "What Is the Point of Equality?"

33 See Bettcher, "Evil Deceivers and Make-Believers."

condition as well, so I do not spend much time on the condition here. It might be possible, however, for some views to meet the disrespect condition without necessarily meeting the disempowerment condition. One could perhaps believe that trans people should, politically speaking, have their right to self-determination granted but still hold them in contempt.³⁴ Other views that simply ignore the existence of trans people may also violate the disrespect condition. For example, expressing the view that only women can get pregnant might not justify limiting trans people's access to maternity wards but still ignores trans men and nonbinary people altogether, which is plausibly a form of disrespect (even if unintentional).

2.2. *How Are Trans People Socially Subordinated?*

Legitimizing views that justify the disempowerment or disrespect of trans people uniquely contributes to their subordination. To show how, let me first explain what I mean by subordination. Here, I follow Moreau, who provides four salient conditions that can be helpful in determining whether a social group is unjustly subordinated.³⁵ These conditions are neither necessary nor sufficient but rather serve as signposts that are present in most relevant cases of subordination. For brevity, I merely mention the first three conditions and focus on the fourth, though it should be uncontroversial that trans people successfully meet each.

The four conditions are as follows. First, the relevant social group possesses less social and political power relative to others. Second, the group possesses or is perceived to possess traits that “attract less consideration or greater censure than the corresponding traits of the empowered group.”³⁶ Third, the negatively perceived traits attributed to the social group contribute to stereotypes that rationalize the group's relative lack of power and lack of *de facto* authority.

One distinguishing feature of Moreau's account is her fourth condition, which concerns what she calls *structural accommodations*. A structural accommodation is a policy, practice, or even physical structure that works “by tacitly *accommodating* the needs or interests of one group and overlooking those of others—with the result that the needs of the dominant group come to seem normal and natural, whereas the different needs of the subordinate group come to seem exceptional and even odd.”³⁷ Moreau uses bathrooms as an example

34 As Hernandez notes, examples like these involve a failure of gender-affirming perception (“Gender-Affirmation and Loving Attention”).

35 Moreau, “Discrimination and Subordination.”

36 Moreau, “Discrimination and Subordination,” 137.

37 Moreau, “Discrimination and Subordination,” 132.

of a structural accommodation that disadvantages many trans people. Despite being designed without malintent, for some trans people, having only men's and women's restrooms has contributed to normalizing the binary-only gender schema, "rendering invisible those who do not fit on one side of it or the other."³⁸ When such structural accommodations work in tandem with stereotypes to justify an unequal distribution of power and *de facto* authority, the group in question satisfies Moreau's fourth condition for social subordination.

2.3. *Institutional Legitimacy as a Structural Accommodation*

The institutional legitimacy of transphobic views is constituted by institutional actors or structures upholding the notion that such views are viable candidates for belief. This legitimacy can be classed as a structural accommodation under Moreau's account. In the case of trans people, structural accommodations include practices and policies that tacitly prioritize the interests of cisgender people at the expense of trans people, which then work in tandem with stereotypes to rationalize trans people's continued lack of social and political power. To be clear, I am not arguing that the mere *expression* of a transphobic view necessarily subordinates trans people but rather arguing that the *legitimizing* of such views through institutional means is a structural accommodation that perpetuates subordination.³⁹

The continued legitimizing of transphobic views involves a system of practices and policies that leaves unanswered the question of whether trans people ought to possess social and political status that is equal to that of cisgender people. Such legitimacy tacitly prioritizes cisgender people's interest in having any and all of their expressed beliefs taken seriously over trans people's claims to social and political equality. By inviting advocates of transphobic views to write opinion pieces, speak on the issue at official university events, or sit on panels of experts, institutions advance the interests of cisgender people at the expense of the trans community. Moreau notes that structural accommodations need not be intended to prioritize the needs of a dominant group in order to be harmful. In academia, the legitimizing of transphobic views may be intended not to subordinate trans people but to foster an environment conducive to truth seeking. While truth-seeking intentions are admirable, legitimization in this context accommodates the interests of a dominant group while neglecting the needs of a subordinate group.

38 Moreau, "Discrimination and Subordination," 133.

39 This contrasts with views like Rae Langton's that hold that some speech acts (e.g., pornography) constitute acts of subordination. See Langton, "Speech Acts and Unspeakable Acts." See also discussion below.

One might be skeptical that legitimizing transphobic views could plausibly prioritize the needs of cisgender people over trans people. To draw from a classic example from John Stuart Mill, questioning the veracity of Newtonian physics does not relegate Newtonians to a lower social status.⁴⁰ However, there is an important disanalogy between standard Millian cases and legitimizing views that justify the subordination of a social group. We can draw on Rae Langton's work on pornography to further illuminate this asymmetry.⁴¹ In addition to arguing that pornography *subordinates* women, Langton argues that pornography *silences* women. In a sense, Langton argues that pornography inhibits women's *de facto* authority by making "certain actions—refusal, protest—unspeakable for women in some contexts."⁴² Subsequently, speech that silences "is not free speech," as it allows one form of speech (pornography) only at the expense of another (women's ability to perform certain speech acts).⁴³

A similar argument can be made when it comes to legitimizing views that rationalize the subordination of trans people. To take transphobic views seriously as viable candidates for belief silences trans people by making it more difficult for them to *do things* with words, such as self-identify, protest, and demand.⁴⁴ Why? In cases where transphobic views are taken seriously, trans people's subordination becomes plausibly justified—their lack of social and political power becomes plausibly warranted. To question the validity of trans people's status in such contexts *necessarily* opens a wedge between trans and cisgender people in which some of the former group's political claims plausibly ought to be denied, while the latter's are not up for discussion. As long as such an inquiry remains open about the status of a subordinated group, the claims of the possibly inferior group are not likely to be given the same uptake as the possibly superior group, even in cases where the possibly inferior group is in fact in an equal or superior epistemic position with regard to the truth of a particular proposition. Thus, such inquiries perpetuate a lack of social power and lack of *de facto* authority for the group whose status is in question.

Furthermore, the institutional legitimacy of transphobic views works in conjunction with stereotypes to rationalize trans people's diminished social and political power. For an institution to take seriously questions about whether trans people's claims to gender self-determination are invalid, trans

40 Mill, *On Liberty*.

41 Langton, "Speech Acts and Unspeakable Acts."

42 Langton, "Speech Acts and Unspeakable Acts," 324.

43 Langton, "Speech Acts and Unspeakable Acts," 327.

44 The idea of *doing things with words* (i.e., performing speech acts) comes from J. L. Austin, whom Langton draws on heavily in her arguments.

women are *actually* men (in the politically relevant sense), or trans people are pathologically confused provides evidentiary support to stereotypes that trans people are mentally ill, deceitful, or dangerous to society (in particular, cisgender women and girls).

Consider, for example, a university hosting a panel themed “Should trans women have access to women-only spaces?” that includes someone debating in the negative. Regardless of how respectful the reasons given may appear, the fact that a transphobic view is presented as a viable candidate for belief by the university’s own lights sends a clear signal to audience members that trans women’s political equality with cisgender women is importantly and *justifiably* contestable. For people who already hold prejudicial views about trans people, this contestability validates stereotypes—it presents itself as *further evidence* that transphobic stereotypes are accurate. Such contestability also provides direct evidence that such stereotypes are plausibly true to those who are undecided and looking to make up their minds.⁴⁵

One might object that while the imagined debate *may* rationalize stereotypes, the aforementioned connection between institutional legitimacy and stereotypes is not *necessary*. However, the link between institutional legitimacy and the rationalization of stereotypes does not need to be airtight on Moreau’s account—that such stereotypes are given justificatory backing through the debate is sufficient for the legitimacy to count as structural accommodation.

There is thus a clear link between the institutional legitimacy of transphobic views and the rationalization of trans people’s diminished social and political status. For these reasons, the institutional legitimacy of transphobic views structurally perpetuates the subordination of trans people. In order to successfully redress such subordination, we must redress the conditions that enable it to exist in perpetuity. One such condition is the institutional legitimacy of transphobic views. While some transphobic structural accommodations may be difficult or even impossible to address, institutionally delegitimizing transphobic views is in fact a possibility in many institutional contexts.

3. TRANSPHOBIC VIEWS, INSTITUTIONAL NORMS, AND ACADEMIC FREEDOM

Having addressed the first two premises of my argument, let me turn to the third: academic institutions can delegitimize transphobic views, at least in many contexts, in a manner that is not excessively burdensome. More specifically, I

45 Thus, having “balanced representation” (which is sometimes called *co-platforming*) still perpetuates the subordination of trans people because such balance implies transphobic views are on par with nontransphobic ones from the perspective of the curator of representation. For more, see Dembroff, Kukla, and Stryker, “Retraction Statement.”

argue that delegitimization aligns with widely endorsed institutional norms (section 3.1) and a robust notion of academic freedom (section 3.2). Along the way, I address the objection that encouraging *more* legitimized trans-affirming speech rather than *less* legitimized transphobic speech constitutes a viable alternative to delegitimization.

3.1. *Transphobic Views and Institutional Norms*

Delegitimizing transphobic views aligns with institutional norms and commitments that most academics, including many advocates for transphobic views, accept. Within an institution, a set of normative assumptions is required for the institution to achieve certain goals related to its flourishing (or even basic functioning). For example, in a racially integrated university, the premise that Black people are morally, politically, and epistemically equal to white people and deserve to be treated as such is taken for granted. Inquiry is closed with regard to the truth of this premise. Racial equality has become a *normative starting assumption*, one that has been deemed both true and necessary to successfully operate a racially integrated institution, with empirical and philosophical backing to justify such a decision. Such assumptions consequently commit the university to a normative ideal of racial equality, which carries with it various obligations regarding how to develop and enact institutional policies and procedures.

A university's set of normative starting assumptions and commitments is commonly thought to be compatible with its epistemic aims, even if a particular normative assumption is directed towards another goal such as racial integration. Robert Mark Simpson and Amia Srinivasan refer to these assumptions as "axioms that define the field."⁴⁶ While Simpson and Srinivasan refer to disciplinary subfields within academia, it is plausible that at least some premises (e.g., racial equality) are axiomatic from the perspective of not only disciplines like philosophy and sociology but also entire universities and other institutions.

Of course, many universities fail to live up to their stated normative commitments. Many universities vocally disavow white supremacy and express a normative commitment to racial equality. In such cases, I take it that the most blatantly white supremacist views, such as the view that people of European heritage are more intelligent than their non-European peers, are in fact institutionally illegitimate. However, within the purview of such institutions, there often exist practices and policies that nevertheless perpetuate white supremacist thinking, such as dismissing or otherwise undervaluing Indigenous or African scholarship. There might also be disagreement about whether a particular

46 Simpson and Srinivasan, "No Platforming," 201.

practice or policy indeed shirks a stated normative commitment. Thus, there can exist plenty of controversy surrounding an institution's normative starting assumptions, including the actions an institution must take to live up to the commitments that such assumptions entail.

However, what I take to be widely endorsed is an institution's making *some* normative starting assumptions and commitments, which in turn become axiomatic. We typically agree that there are contexts in which these assumptions and commitments are perfectly acceptable—namely, when these assumptions and commitments are needed to further a legitimate goal. Even if universities might not in practice behave in ways that perfectly align with their normative commitment to racial equality, we typically agree that the goal of racial integration justifies making certain normative claims about racial equality that are axiomatic from the perspective of the institution. In other words, *were* an institution to completely delegitimize white supremacist views, we would not view this as problematic.

To delegitimize transphobic views, then, is to place the validity of trans people's basic moral, epistemic, and political claims to gender self-determination within an institution's set of normative starting assumptions. Historically, whenever a new normative claim begins to make its way into this set, much resistance arises from those who do not stand to benefit from the addition. This resistance has occurred when considering whether to affirm the equality of cisgender men and women, Black and white people, and nonheterosexual and heterosexual people. Despite previous resistance, many institutions now consider assumptions of equality between these groups to be axiomatic. Such institutions have committed themselves to delegitimizing views that challenge these starting assumptions, even if they do not always perfectly follow through with this commitment. Views that explicitly contradict an institution's normative starting assumptions are not considered viable candidates for belief. If someone presents a valid argument that concludes or implies that women are morally, politically, or epistemically inferior to men, institutional actors assume the argument is unsound, and most do not bat an eye when they assume such unsoundness without carefully analyzing each premise.

So, there exists a set of normative starting assumptions that are required for an institution to meet certain legitimate goals. Such assumptions carry with them a normative commitment to deny legitimacy to certain views. Thus, there is an existing, commonly endorsed practice of institutions establishing a normative commitment to delegitimize certain views. This practice can be extended to include transphobic views. As Michelle Moody-Adams notes, the shared understandings of a particular community of academic inquiry "involve the notion that some ways of arguing, and some points of view, are simply not

worthy of recognition within the community of inquiry.”⁴⁷ Ultimately, then, the current normative commitments of an institution, along with the fact that the institutional legitimacy of transphobic views perpetuates the subordination of trans people, together justify the integration of trans-affirming normative premises into that institution’s set of starting assumptions.

3.2. *Institutional Legitimacy and Academic Freedom*

Delegitimizing transphobic views also aligns with a robust notion of academic freedom. It has been convincingly argued that academic freedom is importantly distinct from free speech broadly construed.⁴⁸ Theories about the value of free speech focus on the role of discourse in society and are often divided into three broad camps, which focus on how free expression underwrites truth-seeking processes, autonomy, and democratic values, respectively.⁴⁹ While theories of free speech inform approaches to academic freedom, the latter focus on a narrower set of protections with regard to academic teaching and research that protects academics from certain forms of political, administrative, or social interference.⁵⁰ Importantly, while the protections constitutive of academic freedom are broad, universities also give control to disciplinary experts in guiding various “communicative norms and practices” within communities of inquiry, creating an asymmetry between the scope of academic freedom and that of free speech.⁵¹ As Simpson notes, certain content-based regulations of speech are actually “integral to the form of intellectual work that the academic performs.”⁵² For an institution to consider some views to be unviable

47 Moody-Adams, “What’s So Special About Academic Freedom?” 106.

48 Simpson and Srinivasan, “No Platforming”; Moody-Adams, “What’s So Special About Academic Freedom?”; Van Alstyne, “The Specific Theory of Academic Freedom and the General Issue of Civil Liberty”; Simpson, “The Relation Between Academic Freedom and Free Speech”; Post, *Democracy, Expertise, and Academic Freedom*; and Fish, *Versions of Academic Freedom*.

49 For more on the relationship between free speech and truth seeking, see Mill, *On Liberty*; Brink, “Millian Principles, Freedom of Expression, and Hate Speech”; and Schauer, “Social Epistemology, Holocaust Denial, and the Post-Millian Calculus. For more on the relationship between free speech and autonomy, see Scanlon, “A Theory of Freedom of Expression”; Strauss, “Persuasion, Autonomy, and Freedom of Expression”; and Brison, “The Autonomy Defense of Free Speech.” For more on the relationship between democracy and free speech, see Meiklejohn, *Political Freedom*; Weinstein, “Participatory Democracy as the Central Value of American Free Speech Doctrine”; and Sunstein, *Democracy and the Problem of Free Speech*.

50 Simpson and Srinivasan, “No Platforming”; and Van Alstyne, “The Specific Theory of Academic Freedom and the General Issue of Civil Liberty.”

51 Simpson and Srinivasan, “No Platforming,” 196.

52 Simpson, “The Relation Between Academic Freedom and Free Speech,” 292.

candidates for belief is not only common practice, then, but inescapable given the nature of academic research. Some papers are not published because their content is considered unserious. Certain views are excluded from syllabi because they do not warrant serious consideration. In fact, plenty of views are institutionally illegitimate because of what Moody-Adams refers to as *practices of exclusion*, which “play a fundamental role in the structures that underwrite academic freedom.”⁵³

Those who seek to protect the promulgation of transphobic views in academic spaces frequently conflate a view’s institutional illegitimacy with a stifling of academic freedom even though these phenomena are not identical. I doubt, for example, that such people believe *all* research programs make a professor worthy of tenure, and *all* journal submissions are worthy of being published. As Simpson and Srinivasan emphatically note,

In the public square we tolerate the speech of flat-earth cranks, shills paid to undermine climate science, and revisionist historians who espouse conspiratorial misreadings of the evidence. As long as they don’t harass anyone we let them say their piece. But such people aren’t owed an opportunity to teach History 101 or publish in scientific journals, any more than they are owed a platform to address parliament or a corporate board meeting. It is permissible for disciplinary gatekeepers to exclude cranks and shills from valuable communicative platforms in academic contexts because effective teaching and research requires that communicative privileges be given to some and not others, based on people’s disciplinary competence.⁵⁴

Institutions need not deny people the ability to express their illegitimate views in certain contexts, including any and all noninstitutional ones. However, academic freedom does not entail that professors, speakers, and students are entitled to have every one of their expressed views taken seriously as viable candidates for belief, especially in contexts that imply institutional endorsement. Simpson and Srinivasan take things even further and argue that academics have a *responsibility* to determine which views should not be taken seriously as viable candidates for belief in service of promoting disciplinary knowledge.

Neither Simpson and Srinivasan nor I deny that a consequence of delegitimization is that some views are given less attention than others, and thus some instances of speech are stifled in the sense that they are less likely to spread and be adopted—that is often the point. This kind of stifling is ubiquitous and

53 Moody-Adams, “What’s So Special About Academic Freedom?” 106.

54 Simpson and Srinivasan, “No Platforming,” 195–96.

arguably necessary when it comes to academic work, even if such stifling is inappropriate in other domains of expression. A white supremacist's speech might be stifled in this weak sense when their article is not published in a philosophy journal, but their academic freedom is not, nor are their basic free speech rights infringed upon. Academics are entitled to effectively stifle white supremacist speech in contexts where institutional endorsement is implied, even if we are not entitled to stifle such speech in the public square. Thus, the distinction between academic freedom and free speech plays a crucial role in justifying the institutional delegitimization of certain views. Stifling academic speech by rejecting an article is not the same kind of regulation as legally suppressing or restricting such speech, the latter of which plausibly requires more justification than the former.

One might object in the following way. Perhaps not all journal articles should be published, and not all candidates should be hired. But what ought to determine our attitudes towards academic speech should have nothing to do with political or moral matters (including redressing subordination). Allowing reasons of subordination to influence whether a view is delegitimized, according to my objector, does not promote disciplinary knowledge. Worse, such permissiveness might even constitute a form of epistemic pollution, compromising the truth-seeking aims of academia. Rather, my objector might argue, views should be evaluated solely on their merits and other disciplinary standards like relevance, methodology, and clarity. I refer to my objector's position as *moral/political weightlessness*—moral and political reasons must not be given any weight whatsoever when evaluating academic speech.

The strongest defense of moral/political weightlessness leans on epistemic defenses of free speech, in particular those found in Mill's *On Liberty*.⁵⁵ Mill defends a robust model of freedom of expression, including the expression of views that we consider certainly false and patently reprehensible. Should either the government (through laws) or the public (through social stigma and sanction) persecute those who express unpopular opinions (including transphobic ones), a variety of deleterious epistemic, moral, and political effects will occur. Further, even though academic freedom is not identical to free speech, Frederick Schauer argues that given the uniquely epistemic aims of academic inquiry and research (which Schauer distinguishes from universities more broadly),

55 Of course, Mill does not present epistemic arguments for free speech only, and epistemic approaches are not the only ones around. However, I take other approaches to free speech, such as those based on autonomy and democracy, to more easily support delegitimization. So I focus here on views that give special importance to safeguarding truth-seeking processes, given that they pose the strongest challenge to delegitimization.

Mill's arguments lend support to relegating nonepistemic considerations "to a decidedly secondary status" in such contexts.⁵⁶

A supporter of moral/political weightlessness, drawing upon Mill, will likely say that if transphobic views are indeed unviable candidates for belief, then we should not be afraid of their being discussed in any context at any time, including those that imply institutional endorsement. The truth will eventually win out, they say, and free speech will have guided the way. Further, they say, it is to trans people's benefit to allow transphobic views to be published and taken seriously because without such vigorous debate, citizens might not truly appreciate the validity of trans people's identities.⁵⁷

This leads us back to the previously mentioned objection that institutions need not delegitimize transphobic views in order to stop perpetuating the subordination of trans people. For many Millians, the best solution to addressing transphobic speech's subordinating potential is not less speech but *more* speech. As the response goes, more transphobic speech will lead to more of what has been termed *counterspeech* from trans-affirming people, and the ensuing dialectic will end trans people's subordination that much faster.⁵⁸ If transphobic views are delegitimized, then the strongest versions of these views may go unexpressed, since academics would not be able to get them published in serious journals. Trans-affirming scholars would then be unable to successfully debunk the best arguments possible (i.e., produce the strongest counterspeech), which could have persuaded many people that transphobic views are false. So the consequences of delegitimization might actually be *worse* in the long run, even from the perspective of redressing the subordination of trans people. At the very least, encouraging counterspeech might be thought to provide a sufficient institutional alternative to delegitimization that could ameliorate trans people's subordination. For my objector, then, even if there exist reasons to institutionally delegitimize transphobic views, such reasons are reliably outweighed by a panoply of Millian reasons, which ultimately speak in favor of moral/political weightlessness.

56 Schauer, "Social Epistemology, Holocaust Denial, and the Post-Millian Calculus," 140.

57 Justice Antonin Scalia makes a similar point in his dissenting opinion in the US Supreme Court Case *Obergefell v. Hodges* 576 U.S. 644 (2015), arguing that the majority opinion granting same-sex couples the right to marry denied the American public the opportunity to be persuaded through vigorous debate.

58 This perspective is exemplified in Justice Louis Brandeis's opinion in the US Supreme Court Case *Whitney v. California* 274 U.S. 357 (1927). For more on counterspeech, see McGowan, "Responding to Harmful Speech"; Bowman and Gelber, "Responding to Hate Speech"; and Tirrell, "Toxic Misogyny and the Limits of Counterspeech."

I offer three reasons to reject moral/political weightlessness, of which the latter two engage directly with Mill. First, while I do not claim to provide an exact test for which moral or political reasons plausibly ought to be allowed into our deliberations about academic speech, at least *some* such reasons are generally already considered fair game for justifying certain content-based regulations of speech. For example, almost all sets of disciplinary norms and practices include normative expectations surrounding respect and professionalism. Even if justified epistemically, norms of respect and civility in academic conduct are undeniably moral. We expect our fellow academics to avoid personal attacks, the divulging of sensitive personal information, and objectionably dismissive tones when performing academic work, even when doing so does not make a difference to the plausibility, novelty, or clarity of a given argument.

Philosopher Stewart Cohen, for example, resigned as editor-in-chief of the journal *Philosophical Studies* because an article by trans philosopher Robin Dembroff contained perceived “unprofessional personal attacks” against philosopher Alex Byrne, who advocates for (on my account) transphobic views.⁵⁹ The “attacks” presumably appeared to violate moral norms that Cohen considered integral to the discipline of philosophy. Similarly, gender-critical feminists objected to the inclusion of the word ‘TERF’ (Trans-Exclusionary Radical Feminist) in an article for the journal *Philosophy and Phenomenological Research*, arguing that the term “denigrate[s] and dismiss[es]” those who advocate for trans-exclusionary policies.⁶⁰ Clearly, then, even those who are not trans affirming believe that certain norms of respect ought to be enforced, and those norms can even warrant public denouncement, calls for apology, or social sanctions for those who violate them.

Enforcing norms of respect falls squarely within the disciplinary control that academics possess over certain instances of academic speech. In other words, such enforcement does not violate standard approaches to academic freedom. Enforcing such norms is generally considered a good thing, even though such enforcement may lead to fewer papers with novel, interesting contributions being published in the style in which some researchers prefer.

To delegitimize transphobic views is arguably a way of enforcing a norm of respect. Trans people are subordinated and disrespected, and they should not receive such treatment. The legitimacy of transphobic views perpetuates their subordination and justifies their being disrespected. Thus, enforcing the delegitimization of transphobic views is a way of enforcing norms of respect

59 The article in question was Dembroff, “Escaping the Natural Attitude About Gender.” For more on the resignation, see Weinberg, “A Resignation at Philosophical Studies and a Reply from the Editors (Updated with Comments from Cohen, Dembroff, Byrne).”

60 Flaherty, “‘TERF’ War.”

for trans people. Those who object to enforcing norms of respect in this case but agree that *some* norms of respect can be considered when engaging in regulatory deliberation must provide a principled reason as to why *their* favored norms of respect are admissible while norms of respect for trans people are not.

The advocate for moral/political weightlessness may push back by distinguishing between *procedure* and *substance*. Perhaps norms of respect apply to only *how* we argue, not *what* we argue. Word choices might be regulated, but the premises they represent cannot be. Yet certain premises are *necessarily* disrespectful, regardless of how politely phrased. Claiming that white people are most likely just the slightest bit better than members of other races (though those people are great, too) still advances a white supremacist view. To advance a white supremacist view disrespects people who are not white. Additionally, in the case of Cohen, the apparent disrespect that led to his resignation was not specific word choice but Dembroff's decision to call into question Byrne's potential political motivations and critique Byrne's lack of engagement with relevant academic literature. In doing so, Dembroff explicitly employs a suggestion Byrne himself makes in his original paper—to urge caution when someone is personally invested in the truth of a particular claim and then testifies that the claim is true.⁶¹ These argumentative moves play a substantive role in the paper and are not only a matter of *how* Dembroff is arguing. I do not think that Dembroff is actually disrespectful for making the moves they make. Rather, I use this example to illustrate how, at least in some cases, it is *substantive content* that is arguably disrespectful according to disciplinary norms, irrespective of how such content is expressed. Enforcing a norm of respect, then, might warrant a content-based restriction of speech.

Second, delegitimizing transphobic views can plausibly be thought to *advance* disciplinary knowledge and promote a healthier epistemic environment, despite accusations of the contrary. As argued above, there is a disanalogy between questioning the axioms of Newtonian physics and questioning the validity of trans people's political claims to gender self-determination. Normative cases that question the moral, epistemic, or political status of subordinated groups are unique because truth seeking *itself* is compromised when such inquiries are open. Regardless of whether a particular transphobic view implies that trans people as a group are less credible, part of what it is to be subordinated is to have less social power and less *de facto* authority. Thus, as long as such a power differential remains, trans people's ability to have any of their views taken as seriously as cisgender people's in relevant social situations

61 Dembroff, "Escaping the Natural Attitude About Gender."

is undermined. They are, to use Langton's word, silenced.⁶² This undermining of credibility subsequently compromises an institution's broader epistemic environment, as it "presents a conflict within liberty itself," between the right to speak (and be taken seriously) of the silencing group and the silenced.⁶³

David Brink makes a similar argument regarding the effects of hate speech from a Millian point of view—the victims of hate speech are effectively silenced when such speech "contributes to a hostile environment that undermines the culture of mutual respect necessary for effective expression and fair consideration of diverse points of view."⁶⁴ As Brink argues, Mill's defense of free expression is based on the immense value (including epistemic value) of one's being able to exercise and develop one's various *deliberative capacities*. Policies regulating hate speech can plausibly be considered a way of safeguarding such capacities and, as such, carve out "a well-motivated exception to the usual prohibition on content-specific regulation of speech."⁶⁵ While I neither argue that all transphobic speech is hate speech nor advocate for the legal regulation of transphobic views, institutionally delegitimizing transphobic views can be similarly viewed as a way of protecting the deliberative capacities of an institution's members, including those who are not transgender. Such protection carries distinctly epistemic benefits without undermining anyone's fundamental rights.

There is thus an important sense in which advocates for moral/political weightlessness, who often claim to be paragons of unadulterated truth seeking, protect a narrow conception of academic freedom while undermining many valuable epistemic goods, at least within our nonideal world. As long as an agent's social identity objectionably undercuts their epistemic authority, dominant social groups remain one step ahead in the supposed free marketplace of ideas. The marketplace, then, is free only for those whose epistemic status is not in question and for those whose speech acts are reliably given uptake.⁶⁶

These considerations cast doubt on the Millian claim that more speech is reliably better than less, considering the aims of *either* redressing subordination *or* truth-seeking. As Mary Kate McGowan argues, when certain forms of speech epistemically disable members of marginalized groups, it is far from

62 See the argument in section 2.3 above.

63 Langton, "Speech Acts and Unspeakable Acts," 329.

64 Brink, "Millian Principles, Freedom of Expression, and Hate Speech," 141.

65 Brink, "Millian Principles, Freedom of Expression, and Hate Speech," 142.

66 For more thorough critiques of the epistemic problems with a free marketplace of ideas, see Goldman and Cox, "Speech, Truth, and the Free Market for Ideas"; and Ingber, "The Marketplace of Ideas."

clear that the group's attempts at counterspeech are even remotely likely to succeed.⁶⁷ Simpson builds on McGowan's earlier work to argue that it is far easier to insert discriminatory speech into a debate than to remove it because it is easier to make something salient in speech than it is to make it unsalient.⁶⁸ This concern is especially relevant when considering speech endorsed by academic journals or universities, given the authority of such venues to make speech salient in a lasting way.⁶⁹ So long as transphobic views remain legitimized, then, merely encouraging counterspeech is liable to be an unreliable strategy at best and counterproductive at worst. Delegitimization, on the other hand, is a viable strategy when it comes to both protecting an institution's deliberative capacities and ameliorating the subordination of trans people.

Furthermore, we do not need to legitimize transphobic views in order to reap the epistemic and political benefits of discussing them, even in journal articles. Many people continue to hold white supremacist views, for example, but the solution to changing such people's minds does not lie in publishing more papers advocating for such views in well-respected journals. Counterspeech can be produced that discusses (and debunks) discriminatory views without thereby considering such views to be viable candidates for belief. It is true that such arguments come from those who believe they are false, which might make the arguments less passionate and carefully crafted. It may also be the case that those of us in the business of academic publishing ought to be more sensitive to the public relevance of journal submissions, since otherwise articles combatting discriminatory views might be seen as uninteresting or obvious from a disciplinary perspective. However, we can be sensitive to such matters and adjust accordingly without thereby endorsing discriminatory views as viable candidates for belief. So one can embrace the idea that generally, more arguments ought to be published that *discuss* transphobic views while rejecting the notion that accomplishing this goal requires institutionally legitimizing such views.

67 McGowan argues that the "more speech" response, as she calls it, generally fails to accurately depict how language is used in many of the contexts in which counterspeech is claimed to be the optimal solution ("Responding to Harmful Speech").

68 Simpson, "Un-Ringing the Bell."

69 There is also an empirical question about the efficacy of counterspeech, of which there is simply insufficient high-quality scholarship with which to make any conclusions (Cepolaro, Lepoutre, and Simpson, "Counterspeech"). If robust empirical evidence were to arise that counterspeech is highly effective at counteracting transphobic attitudes, then perhaps this could weaken the obligation for some institutions to delegitimize transphobic views. Without such evidence, however, I remain highly skeptical of counterspeech as a genuine solution for either redressing subordination or cultivating an epistemically just environment.

In order to cultivate a more epistemically healthy environment, then, institutions ought to create and sustain an epistemically *just* environment, which requires eliminating the association of arbitrary traits like race or gender identity with reduced epistemic credibility (i.e., eliminating epistemic injustice).⁷⁰ An epistemically just environment is significantly harder (if not impossible) to cultivate while transphobic views are institutionally legitimate, even with the production of counterspeech. There is thus a higher-order epistemic reason for institutional actors to allow considerations of subordination to impact whether transphobic views are ascribed legitimacy. Additionally, we can soundly reject the plausibility that counterspeech serves as a viable alternative to delegitimization, completing the defense of my first premise: if an institution does not delegitimize transphobic views, then it perpetuates the subordination of trans people.

The third reason to reject moral/political weightlessness is that many contemporary forms of academic speech arguably cross a line from *speech* into *conduct*, opening such actions up to limits that Mill himself establishes in the fourth chapter of *On Liberty*. According to Mill, once “a person’s *conduct* affects prejudicially the interests of others, society has jurisdiction over it, and the question whether the general welfare will or will not be promoted by interfering with it, becomes open to discussion.”⁷¹ Publishing a journal article, hiring a job candidate, and inviting a speaker are all actions that, while involving speech, cross a threshold from allowing public-square soapboxing to explicit acts of endorsement. Officially endorsing a view as legitimate (e.g., publishing the view in an academic journal) is notably distinct from merely providing a platform for the view to be expressed (e.g., allowing the view to be published on social media). Acts of endorsement are plausibly conduct of the sort that Mill mentions, which implies that delegitimizing transphobic views does not fundamentally undermine anyone’s basic rights to free speech that Mill himself advocates for.

So, if the institutional legitimacy of transphobic views prejudicially affects the basic interests of trans people, while the institutional delegitimization of such views does not undermine anyone’s basic rights to free speech, then on Millian reasoning, institutional actors are in fact warranted in acting upon such moral and political reasons in academia. Considering academic conduct to be more than traditional speech thus speaks against moral/political weightlessness on Millian grounds.

70 See Fricker, *Epistemic Injustice*; Anderson, “Epistemic Justice as a Virtue of Social Institutions”; and Langton, “Speech Acts and Unspeakable Acts.”

71 Mill, *On Liberty*, 127.

The reasons above favor an approach to academic freedom that allows reasons of subordination to be considered when making decisions about academic speech. A trans-inclusive approach to academic freedom is robust and does not involve undue restrictions of freedom by the standards and practices established by most academics with relevant expertise.

4. CONCLUSION

Transphobic views, according to my account, are those that express or imply that being transgender either (i) justifies being denied power and *de facto* authority on par with that of cisgender people or (ii) warrants greater censure or lesser consideration (i.e., disrespect) than being cisgender. Institutions ought to stop taking such views seriously as viable candidates for belief for three main reasons. First, unless they delegitimize transphobic views, academic institutions structurally perpetuate the subordination of trans people. The institutional legitimacy of transphobic views upholds the subordination of trans people by acting as a structural accommodation that works in tandem with stereotypes to rationalize trans people's lack of both power and *de facto* authority. Second, institutions ought to stop perpetuating subordination through their own behavior when stopping is not too burdensome, i.e., it would not compromise the integrity of the institution nor violate norms of academic freedom. Third, institutional delegitimization aligns with both common institutional practices and a robust notion of academic freedom, though it does require abandoning an approach to the latter that demands moral and political reasons be given no weight whatsoever in determining which speech to take seriously. Ultimately, then, the *pro tanto* obligation to delegitimize transphobic views generally outweighs reasons for neutrality. I have sketched a picture as to what delegitimization might look like in practice, while leaving many questions unanswered regarding exactly how this process is best achieved.

It is worth emphasizing the *pro tanto* nature of the obligation I have defended. While delegitimization is not always excessively burdensome, in some cases it might be. Warranted delegitimization should not compromise the integrity of an institution, and in some cases delegitimization could risk doing exactly that, making the obligation to delegitimize defeasible. Take journals that publish empirical views. The integrity of a social science journal may be at risk were it to reject outright any empirical result that lent justificatory support to transphobic policies, such as a high-quality study that finds trans identity to be associated with criminality. As such, it is plausible that empirical journals ought to evaluate findings based exclusively on the quality of methodology and other disciplinary factors in order to maintain social trust and scientific integrity.

There are also important logistical challenges to consider when contemplating how exactly to delegitimize a view, which might speak in favor of delegitimizing with notable caution. For example, there might be strategic incentives from certain parties to push for more views being considered transphobic rather than less, which could lead to unjustifiable censorship. This is a serious worry, as institutions should actively prevent a view's being suppressed merely because certain people in power disagree with it. Proponents of delegitimization must take such challenges into account.

However, there are reasons to suspect that such logistical challenges can be overcome. As I have argued above, institutions already delegitimize views whether we like it or not, often rather subtly. To explicitly delegitimize a view as I have suggested allows for similarly explicit *contestation* when we suspect someone is merely trying to suppress unpopular speech. Decision makers in turn must be prepared to justify their categorizing a view as transphobic with genuine reasons relating to the view's connection to subordination. In cases where such reasons cannot be presented, they may be unjustified in refusing the view uptake. While there may be reasons to delegitimize views, there are also reasons to be cautious when employing this tool and to take corrective action when we suspect it is being abused. Avenues for contesting a view's status as transphobic must exist, then, to maintain a sufficient level of transparency and accountability for institutional decision makers.

It is also worth emphasizing that in arguing for an institutional obligation to delegitimize transphobic views, I do not mean to dismiss the work that trans people themselves have done and will continue to do to foster social equality.⁷² While institutions have a valuable role to play in this liberatory project, trans people themselves deserve to influence and guide these developments. As such, any institutional project of delegitimization should be taken up in a collaborative and deferential manner with and for trans communities, empowering trans people and never behaving in a paternalistic or otherwise condescending manner.

Delegitimization as a strategy to redress subordination should be considered viable not only in the case of transphobic views. Something like delegitimization has already occurred when it comes to overt white supremacy and extreme forms of misogyny. Most universities simply do not consider such views to be viable candidates for belief. Delegitimization may likewise prove to be a useful tool in fostering social equality for other groups. In fact, if my arguments are successful, universities (and other institutions) may have strong reasons to delegitimize *all* views that justify the disempowerment or disrespect

72 See Zurn, "How We Make Each Other."

of any unjustly subordinated group, at least when the legitimacy of such views works in tandem with stereotypes to rationalize the group's lack of social and political power. Such an implication is quite broad.

Nevertheless, as with the case of transphobic views, I urge caution when embarking on any delegitimizing project. Given the second premise of my argument, we ought to delegitimize only when doing so does not constitute an excessive burden. Additionally, we should take care to ensure that we delegitimize only those views that actually perpetuate subordination. We should perhaps risk delegitimizing too few views rather than too many while ensuring sufficient transparency with regard to institutional decision-making. With proper care, however, delegitimization can prove a valuable strategy that sufficiently honors the imperative of social equality alongside the value of academic freedom.⁷³

University of North Carolina at Chapel Hill
lmitchell@unc.edu

REFERENCES

- Anderson, Elizabeth S. "Epistemic Justice as a Virtue of Social Institutions." *Social Epistemology* 26, no. 2 (2012): 163–77.
- . "What Is the Point of Equality?" *Ethics* 109, no. 2 (1999): 287–337.
- Andler, Matthew Salett. "Gender Identity and Exclusion: A Reply to Jenkins." *Ethics* 127, no. 4 (2017): 883–95.
- Ashley, Florence. "Gender Self-Determination as a Medical Right." *Canadian Medical Association Journal* 196, no. 24 (2024): E833–35.
- . "Homophobia, Conversion Therapy, and Care Models for Trans Youth: Defending the Gender-Affirmative Approach." *Journal of LGBT Youth* 17, no. 4 (2020): 361–83.
- . "'Trans' Is My Gender Modality: A Modest Terminological Proposal." In *Trans Bodies, Trans Selves: A Resource by and for Transgender Communities*, 2nd ed. Edited by Laura Erickson-Schroth. Oxford University Press, 2022.
- . "What Is It Like to Have a Gender Identity?" *Mind* 132, no. 528 (2023): 1053–73.
- Ashley, Florence, Shari Brightly-Brown, and G. Nic Rider. "Beyond the Trans/

73 I am extraordinarily grateful to the many friends, colleagues, and mentors who helped me strengthen this paper over the course of many years. I send special thanks to Cheshire Calhoun, Quill Kukla, Tom Dougherty, Daniel Muñoz, Alex Worsnip, Cassie Herbert, and Raja Halwani for their support, encouragement, and incredibly helpful feedback. Additional thanks to my two anonymous reviewers at *JESP*.

- Cis Binary: Introducing New Terms Will Enrich Gender Research." *Nature* 630, no. 8016 (2024): 293–95.
- Aultman, B. "Cisgender." *Transgender Studies Quarterly* 1, nos. 1–2 (2014): 61–62.
- Bettcher, Talia Mae. "Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion." *Hypatia* 22, no. 3 (2007): 43–65.
- Bowman, Kristine L., and Katharine Gelber. "Responding to Hate Speech: Counterspeech and the University." *Virginia Journal of Social Policy and Law* 28, no. 3 (2021): 248–74.
- Briggle, Adam. "Which Reality? Whose Truth? A Review of Kathleen Stock's *Material Girls: Why Reality Matters for Feminism*." *Social Epistemology Review and Reply Collective* 10, no. 11 (2021): 52–59.
- Brink, David. "Millian Principles, Freedom of Expression, and Hate Speech." *Legal Theory* 7, no. 2 (2001): 119–57.
- Brison, Susan J. "The Autonomy Defense of Free Speech." *Ethics* 108, no. 2 (1998): 312–39.
- Bump, Philip. "Nearly 700 Days Later, Most Republicans Still Believe Trump's Big Lie." *Washington Post*, September 28, 2022.
- Byrne, Alex. "Philosophy's No-Go Zone." *Quillette*, April 17, 2023. <https://quillette.com/2023/04/17/philosophys-no-go-zone/>.
- Calhoun, Cheshire. *Feminism, the Family, and the Politics of the Closet: Lesbian and Gay Displacement*. Oxford University Press, 2000.
- Card, Claudia. "Gay Divorce: Thoughts on the Legal Regulation of Marriage." *Hypatia* 22, no. 1 (2007): 24–38.
- Cepollaro, Bianca, Maxime Lepoutre, and Robert Mark Simpson. "Counterspeech." *Philosophy Compass* 18, no. 1 (2023).
- Clare, Stephanie D. *Nonbinary: A Feminist Autotheory*. Cambridge University Press, 2023.
- Darwin, Helana. "Challenging the Cisgender/Transgender Binary: Nonbinary People and the Transgender Label." *Gender and Society* 34, no. 3 (2020): 357–80.
- Dembroff, Robin. "Escaping the Natural Attitude About Gender." *Philosophical Studies* 178, no. 3 (2021): 983–1003.
- Dembroff, Robin, and Cat Saint-Croix. "'Yep, I'm Gay': Understanding Agential Identity." *Ergo* 6, no. 20 (2019): 571–99.
- Dembroff, Robin, Rebecca (now Quill) Kukla, and Susan Stryker. "Retraction Statement." *iai News*, August 26, 2019. <https://iai.tv/articles/retraction-statement-by-robin-dembroff-rebecca-kukla-and-susan-stryker-auid-1256>.
- Fish, Stanley. *Versions of Academic Freedom: From Professionalism to Revolution*. University of Chicago Press, 2014.
- Flaherty, Colleen. "'TERF' War." *Inside Higher Ed*, August 28, 2018. <https://>

- www.insidehighered.com/news/2018/08/29/philosophers-object-journals-publication-terf-reference-some-feminists-it-really.
- Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007.
- Gilley, Bruce. "The Case for Colonialism." *Academic Questions* 31, no. 2 (2018): 167–85.
- Goldman, Alvin I., and James C. Cox. "Speech, Truth, and the Free Market for Ideas." *Legal Theory* 2, no. 1 (1996): 1–32.
- Goodykoontz, Bill. "Why Trump's Fake Victory Speech Was the Worst and Most Dangerous TV of Election Night 2020." *Arizona Republic*, November 4, 2020. <https://www.azcentral.com/story/entertainment/media/2020/11/04/election-2020-trump-speech/6159975002/>.
- Hernandez, E. M. "Gender-Affirmation and Loving Attention." *Hypatia* 36, no. 4 (2021): 619–35.
- Hernandez, E. M., and Rowan Bell. "Much Ado About Nothing: Unmotivating 'Gender Identity.'" *Ergo* (forthcoming).
- Ingber, Stanley. "The Marketplace of Ideas: A Legitimizing Myth." *Duke Law Journal* 1984, no. 1 (1984): 1–91.
- Kahn, Chris. "53% of Republicans View Trump as True US President." *Reuters*, May 24, 2021. <https://www.reuters.com/world/us/53-republicans-view-trump-true-us-president-reutersipsos-2021-05-24/>.
- Langton, Rae. "Speech Acts and Unspeakable Acts." *Philosophy and Public Affairs* 22, no. 4 (1993): 293–330.
- Lawford-Smith, Holly. *Gender-Critical Feminism*. Oxford University Press, 2022.
- Levy, Neil. "No-Platforming and Higher-Order Evidence, or Anti-Anti-No-Platforming." *Journal of the American Philosophical Association* 5, no. 4 (2019): 487–502.
- McCarthy, Justin. "US Approval of Interracial Marriage at New High of 94%." *Gallup*, February 21, 2022. <https://news.gallup.com/poll/354638/approval-interracial-marriage-new-high.aspx>.
- McGowan, Mary Kate. "Responding to Harmful Speech: The More Speech Response, Counter Speech, and the Complexity of Language Use." In *Voicing Dissent*, edited by Casey Rebecca Johnson. Routledge, 2018.
- Meiklejohn, Alexander. *Political Freedom: The Constitutional Powers of the People*. Harper and Brothers, 1960.
- Mill, John Stuart. *On Liberty*. Floating Press, 1859/2017.
- Moody-Adams, Michele. "What's So Special About Academic Freedom?" In *Who's Afraid of Academic Freedom?* edited by Akeel Bilgrami and Jonathan R. Cole. Columbia University Press, 2015.

- Moorhead, Joanna. "Kathleen Stock: Taboo Around Gender Identity Has Chilling Effect on Academics." *Guardian*, May 22, 2021.
- Moreau, Sophia. "Discrimination and Subordination." In *Oxford Studies in Political Philosophy*, vol. 5, edited by David Sobel, Peter Vallentyne, and Steven Wall. Oxford University Press, 2019.
- Post, Robert C. *Democracy, Expertise, and Academic Freedom: A First Amendment Jurisprudence for the Modern State*. Yale University Press, 2012.
- Scanlon, Thomas. "A Theory of Freedom of Expression." *Philosophy and Public Affairs* 1, no. 2 (1972): 204–26.
- Schauer, Frederick. "Social Epistemology, Holocaust Denial, and the Post-Millian Calculus." In *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, edited by Michael Herz and Peter Molnar. Cambridge University Press, 2012.
- Serano, Julia. *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity*. Hachette UK, 2016.
- Simpson, Robert Mark. "The Relation Between Academic Freedom and Free Speech." *Ethics* 130, no. 3 (2020): 287–319.
- . "Un-Ringing the Bell: McGowan on Oppressive Speech and the Asymmetric Pliability of Conversations." *Australasian Journal of Philosophy* 91, no. 3 (2013): 555–75.
- Simpson, Robert Mark, and Amia Srinivasan. "No Platforming." In *Academic Freedom, Engaging Philosophy*, edited by Jennifer Lackey. Oxford University Press, 2018.
- Stock, Kathleen. *Material Girls: Why Reality Matters for Feminism*. Hachette UK, 2021.
- . "Response to Christa Peterson's Blog." *Kathleen Stock* (blog). Archived at Wayback Machine, <https://web.archive.org/web/20220815043606/https://kathleenstock.com/response-to-christa-petersons-blog/> (accessed May 13, 2025).
- Strauss, David A. "Persuasion, Autonomy, and Freedom of Expression." *Columbia Law Review* 91, no. 2 (1991): 334–71.
- Sunstein, Cass. *Democracy and the Problem of Free Speech*. Free Press, 1993.
- Tirrell, Lynne. "Toxic Misogyny and the Limits of Counterspeech." *Fordham Law Review* 87, no. 6 (2019): 2433–52.
- Van Alstyne, William. "The Specific Theory of Academic Freedom and the General Issue of Civil Liberty." In *The Concept of Academic Freedom*, edited by Edmund L. Pincoffs. University of Texas Press, 1972.
- Veber, Michael. "The Epistemology of No Platforming: Defending the Defense of Stupid Ideas on University Campuses." *Journal of Controversial Ideas* 1, no. 1 (2021).

- Weinberg, Justin. "A Resignation at *Philosophical Studies* and a Reply from the Editors (Updated with Comments from Cohen, Dembroff, Byrne)." *Daily Nous*, June 12, 2020. <https://dailynous.com/2020/06/12/resignation-philosophical-studies-reply-editors/>.
- Weinstein, James. "Participatory Democracy as the Central Value of American Free Speech Doctrine." *Virginia Law Review* 97, no. 3 (2011): 491–514.
- Zurn, Perry. "Crippling Cis: Rethinking Cisgender Within a Disability Critique." *Journal of Philosophy of Disability* (2024).
- . *How We Make Each Other: Trans Life at the Edge of the University*. Duke University Press, 2024.

GASLIGHTING AND EPISTEMIC COMPETENCE

Cameron Domenico Kirk-Giannini

ONE ATTRACTIVE IDEA about gaslighting is that it can be individuated by the distinctive kind of *dilemma* with which the gaslighter confronts his victim: she must choose between trusting him, in which case she has good reason to doubt her epistemic competence, and rejecting his testimony, in which case she must negotiate the interpersonal consequences of concluding that he is unreliable, manipulative, or perhaps even malicious.

A dilemmatic account of gaslighting along these lines is *anti-intentionalist* in the sense that it holds that a gaslighter need not have any particular intention beyond the intention to communicate *p* to his victim. Anti-intentionalist accounts of gaslighting have a number of appealing features. For example, they can easily account for the intuition that gaslighting occurs in certain paradigm cases. Consider the following case, which the literature has regarded as a clear example of gaslighting:

Skeptical Peers: I moved out of one field of philosophy in grad school due to an overwhelming accumulation of small incidents. . . . When I tried to describe to fellow grad students why I felt ostracized or ignored because of my gender, they would ask for examples. I would provide examples, and they would proceed through each example to “demonstrate” why I had actually misinterpreted or overreacted to what was actually going on.¹

Skeptical Peers contains no information about the intentions of the victim’s peers, and the strength of the intuition that they gaslight her is not diminished by stipulating that they lack the intention to manipulate or undermine her: they are not excused from the charge of gaslighting if they disregard her testimony simply out of perverse contrarianism, for example. Anti-intentionalist accounts of gaslighting may also be easier to apply in practice, since they do not require us to reach firm conclusions about the internal motivations, intentions, or prejudices of speakers before deciding whether they have engaged in gaslighting.

However, anti-intentionalist accounts of gaslighting that center its dilemmatic structure risk overgenerating. Whenever a speaker communicates some

1 This case is from Abramson, “Turning Up the Lights on Gaslighting,” 5.

proposition that is inconsistent with the beliefs of one of her interlocutors, that interlocutor must choose whether to accept the testimony and revise his beliefs or reject the testimony and conclude that the speaker has spoken falsely. So more must be said to explain why everyday cases of disagreement do not constitute gaslighting.

I have explored one approach to this problem in my 2023 article “Dilemmatic Gaslighting,” where I defend a dilemmatic account of gaslighting:

Dilemmatic Gaslighting: For all persons *A*, *B*, and propositions *p*: *A* gaslights *B* with respect to *p* iff (i) *A* intentionally communicates *p* to *B*, (ii) *B* knows (and *A* is in a position to know) that if *p* is true, then *B* has good reason to believe that she lacks basic epistemic competence in some domain *D*, (iii) *A* does not correctly and with knowledge-level doxastic justification believe *p*, and *A* does not correctly and with knowledge-level doxastic justification believe that *B* lacks basic epistemic competence in *D*, and (iv) *B* assigns significant weight to *A*’s testimony.²

My earlier account contains several conditions designed to address worries about overgeneration. In what follows, I focus on the condition that a gaslighter must give his victim reason to doubt her *basic* epistemic competence in some domain.³ I argued that incorporating this condition into Dilemmatic Gaslighting helps it deal with a range of cases that would otherwise be problematic. For example,

It seems to me that it is not possible ... for one paleontologist to gaslight another by suggesting that her considered view about what caused the extinction of the dinosaurs is implausible.⁴

There are some domains in which our beliefs are not plausibly regarded as formed on the basis of any basic epistemic competence. First, there are beliefs about theoretical domains like advanced mathematics, the natural and social sciences, philosophy, and so forth. If you demonstrate that I have made some mistake in a complex calculation involving the physics of lasers, I do not thereby gain a reason to doubt any basic epistemic competence of mine. The same can be said about most areas of philosophy.⁵

2 This case appears in Kirk-Giannini, “Dilemmatic Gaslighting,” 757.

3 The idea that a gaslighter targets the basic competence of his victim is also suggested by Abramson: “The accusations are about the target’s basic rational competence—her ability to get facts right, to deliberate, her basic evaluative competencies and ability to react appropriately” (“Turning up the Lights on Gaslighting,” 8).

4 Kirk-Giannini, “Dilemmatic Gaslighting,” 756.

5 Kirk-Giannini, “Dilemmatic Gaslighting,” 765.

Suppose our colleague Professor Plum gives us a cryptic smirk at the department colloquium. You think he means to indicate that he has a devastating objection to the speaker's theory; I think he means to indicate that he has once again succeeded in pilfering one of the bottles of wine meant for the reception. Even if your belief is correct, my insisting on my wine hypothesis does not call into question any basic epistemic competence of yours—Plum's smirk was, after all, cryptic.⁶

In each of these cases, we have something like the dilemmatic structure present in gaslighting, but the intuitive verdict is that gaslighting does not occur. I explained this datum by (i) requiring gaslighting to target a basic epistemic competence and (ii) maintaining that the disagreements in question do not call into question any basic epistemic competence of the parties involved.

The idea that gaslighting must target a basic epistemic competence is thus *prima facie* attractive from the perspective of an anti-intentionalist, dilemmatic theory of gaslighting. Unfortunately, however, Scott Hill shows that enforcing a connection between gaslighting and basic epistemic competence leads to problems of undergeneration. In particular, there are intuitive examples of gaslighting in which the target is not a basic epistemic competence. For example, Hill presents the following version of *Skeptical Peers*:

Skeptical Peers III: Paula tells her peers that she feels ostracized and ignored in her subfield of philosophy because she is a woman. Paula provides examples to illustrate. She evaluates those examples via her views about complicated statistical inferences, sociological background claims, and philosophical reflection about how women in philosophy are generally treated. Her peers know that she is right. But they dismiss her concerns as being based on a misunderstanding of complicated statistics. They tell her that because she is a woman she is incapable of competently engaging in the kind of advanced statistical reasoning required to understand the examples. They maintain that while women have all basic epistemic competences, they do not have the advanced epistemic competences that are unique to men. Distressed, Paula begins to wonder whether they might be right. And she thinks she might be misunderstanding the complicated statistics and therefore whether she has been discriminated against.⁷

6 Kirk-Giannini, "Dilemmatic Gaslighting," 765.

7 This case is from Hill, "Gaslighting and Peer Disagreement," 644. Hill also presents another case:

Skeptical Peers II: Paula tells her peers that she feels ostracized and ignored in her subfield of philosophy because she is a woman. Paula provides examples to

An account of gaslighting that ties it constitutively to basic epistemic competence predicts that Paula's peers in *Skeptical Peers III* do not gaslight her since they do not give her reason to doubt any basic epistemic competence. But, Hill argues, *Skeptical Peers III* is intuitively a case of gaslighting. And if we hold that it is important to accommodate our intuitions about versions of *Skeptical Peers* where the peers lack any intention to manipulate or undermine, it must also be important to accommodate our intuitions about a version of *Skeptical Peers* where the epistemic competence targeted is not basic. As Hill puts it:

There are no details in the original *Skeptical Peers* about whether what is called into question is the graduate student's knowledge from a position of marginalization or her ability to do complicated statistics or anything else. . . . This suggests that exactly which epistemic competence is called into question is not relevant to our intuitions about whether she is gaslighted.⁸

Cases like *Skeptical Peers III* thus appear to pose a problem for anti-intentionalist accounts of gaslighting that center its dilemmatic structure. To avoid overgeneration, they face pressure to hold that gaslighting must target a basic epistemic competence. But to avoid undergeneration, they face pressure *not* to hold this.

What might a friend of Dilemmatic Gaslighting and related views say in response to cases like *Skeptical Peers III*? My sense is that the intuition that *Skeptical Peers III* is a case of gaslighting is not profitably denied. One option is to fall back on the methodology of conceptual engineering, arguing that the theoretical utility of Dilemmatic Gaslighting renders it resistant to refutation by clever counterexamples like those suggested by Hill: while it is important to accommodate our intuitions about cases, it is not *all*-important. But to respond in this way would be unsatisfying insofar as it would do nothing to explain why

illustrate. When Paula considers the examples, they seem to her to clearly be cases that illustrate discrimination. When her peers consider the cases, they seem to them to clearly not be such cases. Paula forms her belief on the basis of her personal experiences. Paula's peers form their belief on the basis of statistical reasoning about her descriptions of the case. Paula and her peers assign significant weight to each other's testimony. (642)

At first, Hill suggests that Dilemmatic Gaslighting might counterintuitively predict that Paula could be gaslighting her peers in *Skeptical Peers II*. However, he goes on to note that there is a compelling reply to this worry: Paula's peers do not form their belief using any basic epistemic competence, so Paula's challenging that belief cannot call into question any basic epistemic competence of theirs. For this reason, he does not offer *Skeptical Peers II* as an objection to Dilemmatic Gaslighting.

8 Hill, "Gaslighting and Peer Disagreement," 646.

we intuit that gaslighting occurs in cases like Skeptical Peers III. Instead, I suggest that there is a different response that can accommodate cases like Skeptical Peers III without giving up the core commitments of the dilemmatic approach.

The response I recommend is to revise Dilemmatic Gaslighting by removing both occurrences of the word 'basic'. The resulting view, which we might call Dilemmatic Gaslighting*, retains the other important features of my original proposal without restricting gaslighting to basic epistemic competences. In this way, it generates intuitive verdicts about both Skeptical Peers III and Skeptical Peers.

As we have seen, however, the condition that gaslighting must target a basic epistemic competence is designed to do explanatory work in avoiding over-generation problems. So it is incumbent on me to explain how such problems can be avoided without restricting gaslighting to basic epistemic competences.

According to Dilemmatic Gaslighting*, *A* gaslights *B* with respect to *p* only if: if *p* is true, *B* has good reason to believe that she lacks epistemic competence in some domain *D*. Whether this condition is satisfied in any given case depends on the strength of the evidence *p* provides that *B* lacks epistemic competence in *D*, which in turn depends on the identities of *p* and *D*. My suggestion is that reflection on cases involving various communicated propositions and epistemic domains significantly diminishes the force of overgeneration worries. In other words, the problem that the basic epistemic competence condition was meant to solve never existed in the first place.

Say that a communicated proposition *p* is a *strong challenge* to hearer *B*'s epistemic competence in domain *D* just in case (if it is true) *p* is good evidence that *B* lacks epistemic competence in *D*. Then the crucial question for Dilemmatic Gaslighting* is whether our intuitions about whether gaslighting has occurred in a case line up in the right way with the facts about whether the communicated proposition is a strong challenge to the hearer's epistemic competence in that case. I believe our intuitions do line up in this way.

Consider first Skeptical Peers III. In this case, the proposition that constitutes the gaslighting—that women are "incapable of competently engaging in ... advanced statistical reasoning"—*entails* in the context that the victim lacks epistemic competence in the relevant domain (advanced statistics).⁹ This is the strongest possible evidential connection between *p* and the proposition that *B* lacks epistemic competence in *D*, so it is no surprise that we have the intuition that gaslighting has occurred.

9 I remain neutral on the question of whether one's lacking epistemic competence in advanced statistics entails that one also lacks competence in statistics simpliciter.

This situation contrasts with ordinary cases of peer disagreement implicating a nonbasic epistemic competence. Consider my earlier examples: first, beliefs in theoretical domains like physics, philosophy, and paleontology; second, “beliefs which, while they do not belong to theoretical domains, are formed on the basis of evidence which is subtle or otherwise difficult to interpret” like our beliefs about Professor Plum’s cryptic smile.¹⁰

Take the paleontology example. It still seems to me that learning the proposition that one’s considered view about what caused the extinction of the dinosaurs is implausible would not call into question one’s basic epistemic competence in any domain. But now I think a stronger claim is also plausible: this proposition is not a strong challenge to *any* of one’s epistemic competences, basic or otherwise. By their nature, the kinds of advanced epistemic competences on which paleontologists rely in forming their views about what caused the extinction of the dinosaurs can misfire. Even if one is perfectly epistemically competent, one might arrive at the wrong view about the extinction of the dinosaurs by failing to detect some subtle methodological flaw in a research paper or by relying slightly too heavily on one source of evidence—or in any number of other ways. Just as we would not interpret a single error in computing a difficult arithmetic problem using pen and paper as good evidence of arithmetic incompetence (or even incompetence at advanced arithmetic), we would not interpret these kinds of mistakes as good evidence of paleontological incompetence (or even incompetence at advanced paleontology). And the same point applies also to other theoretical domains, including physics and philosophy.¹¹

Similar remarks apply to the cryptic smirk case. Here, however, the explanation for why my insisting on my wine hypothesis does not constitute a strong challenge to your epistemic competence is not that interpreting smirks in general requires difficult cognitive work. The explanation is rather that that *in this particular case*, the evidence is objectively ambiguous or difficult to interpret.

If this line of reasoning is sound, we should expect that a version of Skeptical Peers III in which Paula’s peers fail to provide a strong challenge to Paula’s epistemic competence will generate less of an intuition that gaslighting has occurred. This does seem to be the case: if we modify the case so that the peers’ testimony to Paula is simply that her statistical reasoning must have gone wrong somewhere on this occasion (without further explanation), for example, the

10 Kirk-Giannini, “Dilemmatic Gaslighting,” 765.

11 As well as statistics—which, it is worth noting, provides the present proposal with a way of responding to Hill’s worry that Paula might be gaslighting her peers in Skeptical Peers II: she cannot be gaslighting them because the proposition she communicates does not constitute a strong challenge to their statistical competence.

situation seems better described as one in which they are bullshitting (in the sense that Harry Frankfurt describes) than as one in which they are gaslighting her.¹²

So only in certain special cases does Dilemmatic Gaslighting* predict that disagreements implicating an advanced epistemic competence involve gaslighting: cases in which the communicated proposition constitutes a strong challenge to the hearer's epistemic competence. These are generally cases in which what is communicated goes beyond merely the claim that the addressee has gotten things wrong on a particular occasion.

Interestingly, the same does not seem to be true of basic epistemic competences. If you are looking at a large yellow vase under ideal conditions, and I try to convince you that the vase is in fact blue, then I am gaslighting you. This is plausibly because basic epistemic competences, unlike advanced epistemic competences, do not normally misfire. While even the most competent arithmetician makes errors in complicated calculations from time to time, mistaking a blue vase for a yellow one under ideal conditions even once is good evidence that an individual's perceptual apparatus is not functioning properly. So saying that one has gotten things wrong on some occasion with respect to a basic epistemic competence often constitutes a strong challenge to that competence. Given this difference between basic and nonbasic epistemic competences, Dilemmatic Gaslighting* can explain why the most intuitively forceful cases of gaslighting often involve calling into question an individual's basic epistemic competence in some domain.

The preceding discussion raises an important issue for philosophers interested in gaslighting, whether or not they are attracted to an anti-intentionalist view like Dilemmatic Gaslighting: What necessary and sufficient conditions must be satisfied for a communicated proposition to constitute a strong challenge to a hearer's epistemic competence in a given domain? While I have offered a few generalizations above, I leave a more substantive answer to this question for future research.

Rutgers University–Newark
cdk58@philosophy.rutgers.edu

REFERENCES

Abramson, Kate. "Turning Up the Lights on Gaslighting." *Philosophical Perspec-*

¹² See Frankfurt, *On Bullshit*.

tives 28, no. 1 (2014): 1–30.

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

Hill, Scott. “Gaslighting and Peer Disagreement.” *Journal of Ethics and Social Philosophy* 26, no. 3 (2024): 641–47.

Kirk-Giannini, Cameron Domenico. “Dilemmatic Gaslighting.” *Philosophical Studies* 180, no. 3 (2023): 745–72.

JOURNAL of ETHICS & SOCIAL PHILOSOPHY
<http://www.jesp.org>
ISSN 1559-3061

The *Journal of Ethics and Social Philosophy* (JESP) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge. Articles are typically published under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license, though authors can request a different Creative Commons license if one is required for funding purposes.



Funding for the journal has been made possible through the generous commitment of the Division of Arts and Humanities at New York University Abu Dhabi.

جامعة نيويورك أبوظبي



NYU ABU DHABI