

JOURNAL *of* ETHICS
& SOCIAL PHILOSOPHY

VOLUME XIII · NUMBER 2

May 2018

ARTICLES

- 91 A Tripartite Theory of Love
Sam Shpall
- 125 Practical Reason Not as Such
Kenneth Walden
- 154 In or Out? On Benevolent Absolutisms
in *The Law of Peoples*
Robert Huseby

DISCUSSION

- 179 Shared Intention Is Not Joint Commitment
Matthew Kopec and Seumas Miller

The *Journal of Ethics and Social Philosophy* (ISSN 1559-3061) 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 and is available to all readers under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license. Funding for the journal has been made possible through the generous commitment of the Gould School of Law and the Dornsife College of Letters, Arts, and Sciences at the University of Southern California.

The *Journal of Ethics and Social Philosophy* 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 associate 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.

Editor

Mark Schroeder

Associate Editors

James Dreier

Julia Driver

David Estlund

Andrei Marmor

Discussion Notes Editor

Douglas Portmore

Editorial Board

Elizabeth Anderson	Philip Pettit
David Brink	Gerald Postema
John Broome	Joseph Raz
Joshua Cohen	Henry Richardson
Jonathan Dancy	Thomas M. Scanlon
John Finnis	Tamar Schapiro
John Gardner	David Schmidtz
Leslie Green	Russ Shafer-Landau
Karen Jones	Tommie Shelby
Frances Kamm	Sarah Stroud
Will Kymlicka	Valerie Tiberius
Matthew Liao	Peter Vallentyne
Kasper Lippert-Rasmussen	Gary Watson
Elinor Mason	Kit Wellman
Stephen Perry	Susan Wolf

Managing Editor

Susan Wampler

Typesetting

Matthew Silverstein

A TRIPARTITE THEORY OF LOVE

Sam Shpall

ONE PSYCHOLOGICAL CONDITION has an especially deep connection to the felt meaningfulness of life. The purpose of this essay is to explore the nature of this condition, which I think it is appropriate to call *love*.

I begin by more precisely articulating my motivating platitude about love and meaning. Having identified this crucial and underappreciated element of love's psychological profile, I use it to select a varied set of examples that it is natural to think of as paradigms. I observe that almost all prominent philosophical accounts of love are, surprisingly, silent about some of these paradigm cases. The aim is not to show that these theories are misguided, but to suggest that my approach is distinctive and worth exploring.

I then offer a tripartite theory of love's character. Love is *devotion* that renders *vulnerable* and expresses *liking*.¹ The theory is in part a response to the current state of philosophical perplexity about love, where controversy abounds concerning even the most basic questions.² One ambition is to provide the requisite psychological detail to make classic debates about love more tractable. For example, the theory may shed new light on questions about whether love is an emotion, whether love is under our control, whether love for various kinds of nonhuman and nonliving objects is possible, and whether love is rationally assessable. Another ambition is to identify important choices we confront in

- 1 Concerning devotion, the most obvious precursor is Frankfurt, *The Reasons of Love*. On vulnerability, the most ambitious and interesting discussion is Velleman, "Love as a Moral Emotion." Regarding liking, I am influenced by the treatment of friendship in Nehamas, *On Friendship*. I trace many more threads of inspiration in what follows.
- 2 Philosophers disagree about whether love is an emotion (Velleman, "Love as a Moral Emotion"; Brogaard, *On Romantic Love*), a structure of will (Frankfurt, "On Caring," "Autonomy, Necessity, and Love," and *The Reasons of Love*), a form of valuing (Kolodny, "Love as Valuing a Relationship"), or a special mode of perception (Jollimore, *Love's Vision*). They also disagree about how to approach the topic: for example, they disagree about what count as paradigm cases. For a discussion of some of the methodological problems to which this essay responds, see Soble, *The Philosophy of Sex and Love*, e.g., 129.

thinking systematically about love, choices that are sometimes obscured in the vast literature on the topic.

1. THE EXPERIENCE OF MEANING

Here is my motivating thought. More than anything else, love makes human lives seem meaningful to those who live them. It is the most robustly meaning-generating psychological condition with which we are familiar.³

I cannot offer a developed account of the perception or experience of meaning. Few philosophers of love have such accounts.⁴ Yet all of us are familiar with the experience of finding special meaning in, say, a friend's unexpectedly sensitive gesture. For my purposes, it will suffice to make two simple observations that help to distinguish the experience of meaning from experiences with which it might be confused.

First, the experience of meaning is not the same as the perception of objective value. You can find something to be intensely meaningful that you know to be relatively unimportant from the point of view of the universe: consider an old and admittedly over-treasured shirt that you would be devastated to lose, or a much-appreciated visit from a relative who could be spending more time on his medical research. And you can find little or no meaning in something you know to be objectively important: consider an affectless effective altruist, who recognizes the value of her donations even though they do not make a dent in her boundless ennui.⁵

Second, the experience of meaning is not identical to the more common and often trivial experience of desire-satisfaction. (Alternative formulation: it is not the case that all pro-attitudes play the robust meaning-generating role common sense assigns to love.) I am thirsty. When I take a drink in a moment I will not find it particularly meaningful, or meaningful at all. I want to control my back

3 This does not make meaning-generation a necessary condition for love. (Suppose intention is the most behavior-producing psychological condition. It does not follow that behavior-production in any given case is necessary for a state to be an intention.) I am hesitant to endorse the necessity claim about love and meaning, though it may well be true. The robustness claim is a sufficiently useful point of departure.

4 But see Susan Wolf's stimulating reflections in *Meaning in Life and Why It Matters*.

5 For a profoundly distressing example that resembles this, see Rae Langton's brilliant discussion of Maria von Herbert, who finds no meaning in her moral uprightness ("Duty and Desolation"). A similar case is that of J. S. Mill's depression. His narration of this period includes the following wise conjecture: "If I had loved any one sufficiently to make confiding my griefs a necessity, I should not have been in the condition I was" (*Autobiography of John Stuart Mill*, 95).

pain. I take some aspirin and it helps. I value this outcome, but to speak of finding meaning in it sounds forced—indeed, the pain may simply slip away without my noticing. Sally intends to go to work each day, and desires to keep her job, but only for its instrumental benefits. The fact that she does not find it meaningful is exactly why we counsel her to consider a change of career. Ahmed is strongly motivated to pursue casual sex, but he readily acknowledges that he does not find much meaning in it. This explains his simultaneous search for a different kind of intimate relationship.

I cannot make the distinctions between judging things to be valuable, getting things we desire, and finding things meaningful totally precise. I can maintain, however, that the perception of meaning is distinctive, and central to human flourishing. It is largely because we find meaning in life that we continue to live it in the face of hardship. The psychological and moral importance of love is in part explained by its robust connection to this experience of meaning.⁶

One goal of this essay is to develop an account of love that delineates the composition of such a meaning-generating condition. My approach is a broadly functionalist one. I attempt to understand love better by understanding more about this special psychological role. The methodology contrasts with a prevailing (if usually unstated) assumption that analysis of love should begin with paradigm cases, chosen because they exemplify a privileged kind of love object or relationship. I begin with my own set of paradigm cases, selected not for uniformity of object or relationship type, but for exemplification of the meaning-generating role. My cases are chosen to stress the wide variety of sources of meaning in life, and to focus attention on forms of love that I think are obviously interesting but are more frequently discussed by historical than contemporary philosophers.

This methodological point is central. Here is another way of expressing it. There are several fundamental clashes of intuition in sophisticated thinking about love. One such clash concerns which cases are paradigms of the phenomenon. This is a consequential disagreement, since the selection of paradigms determines the contours of our theories. Many thoughtful writers take it to be obvious that the paradigms of love are to be identified on the basis of object kind

6 For some versions of my motivating thought about love and meaning, see Singer, “From *The Nature of Love*,” and Ferracioli, “The State’s Duty to Ensure Children Are Loved,” 8. Having already registered my debt to Wolf (*Meaning in Life and Why It Matters*), I will mention my worry that her conception of meaning, as the intersection of objective value and subjective attraction, is unstable. I cannot defend that claim here. But it is worth emphasizing that we can find things meaningful in ways that do not line up with our considered judgments about value. The observation frames a big question for moral philosophy: To what degree *should* our loving accord with our value judgments?

(e.g., person) or relationship type (e.g., romantic partnership). But the more uniformity our treatments exhibit, the more we should worry about what we are missing.⁷

Having said this, I should make a clarifying note about my aspirations. I am uninterested in policing use of an English term. I am also uninterested in taking an unnecessarily adversarial stance toward the work that has stimulated my interest in these topics. The view of love articulated here is more capacious than the views given by prominent philosophers, but this does not mean that our accounts are incompatible. It is perfectly legitimate for different thinkers to have different theoretical ambitions. I only ask the reader to consider whether it may be *generally* unfortunate that the literature on love is intensely focused on a relatively homogeneous set of examples.

2. PARADIGMS

I now present five cases that give voice to commonsense judgments about the variety of love, and that exemplify the meaning-generating truism that is my point of departure. There are two main purposes to this presentation. I aim, first, to ground the theoretical reflections to come in ordinary, recognizable phenomena. My tripartite theory will unify these examples, and others like them, by articulating the fundamental underlying similarities between their protagonists in more precise detail. Second, the discussion sets up my observations in the next section about the state of the contemporary philosophical literature, which concern the ways in which prominent accounts of love must ignore at least some of what I and many others regard as paradigm cases. In particular, I will highlight the common convention of tailoring philosophical accounts of love to the special case of love for persons, or to special cases of that special case.

My cases thus offer an alternative to some widespread trends of example choice that are perhaps less than fully imaginative, and that enshrine controversial intuitions.⁸ I try to describe them in a neutral way that does not presuppose

7 I will be particularly interested to contrast my approach with the common (exclusive) interest in love for persons. Even those who worry about this general orientation (Smuts, “Normative Reasons for Love, Parts I and II,” 509), or pitch a functionalist story (Jenkins, “Modal Monogamy,” 356), rarely develop such thoughts in any detail. But I will also eschew other commonplace assumptions—e.g., about monogamy (Nozick, “Love’s Bond”) or the necessity of relationships (Kolodny, “Love as Valuing a Relationship”).

8 Some writers express sympathy with my general orientation without exploring its consequences in the same sort of detail. They include Frankfurt, *The Reasons of Love*; Nehamas, *Only a Promise of Happiness*; and Wolf, *The Variety of Values*. Plato and Freud are interesting comparisons, as I note below.

my interpretation of them as paradigms of love. Of course, I invite readers to think of them in this way, and will later help myself to the description. But what matters most is that the examples motivate the coherence and interest of a unified analysis of their protagonists' psychological conditions. If I can convince readers that this analysis is worthwhile and that it has rarely been undertaken, then my main preliminary goals will have been accomplished. I present the cases without interruption before offering further commentary.

World's Best Mom: Max's mom Sherry raised him right. Though there were some turbulent moments between them, especially as they traversed his adolescence, Max recognizes that Sherry consistently acted with compassion, grace, foresight, and great affection. And quite apart from the filial debt he feels, Max also genuinely appreciates Sherry's personality, and enjoys spending time with her. He revels in her successes, helps her when she is in need, and worries about her comfort and happiness as she ages. He does these things not just because he wants to respect what she's done for him, and not just because he enjoys her company and personality, but also because he values her flourishing for its own sake.

Talk to Her: After decades of disappointing searches for a soul mate Nancy decides to embark on motherhood alone. She gives birth to a baby girl she names Alicia, and has never felt so attached to another being or so arrested by the immediacy and depth of her joy. Nancy knows that she would make great sacrifices for Alicia. She is prepared to continue dedicating her life to this fragile creature. However, Nancy's doctors soon discover that Alicia has an untreatable degenerative disease, which will render her brain function catastrophically impaired and take her life within a few years. Nancy is crushed by this revelation. Still, she commits to Alicia with renewed vigor, and spends the happiest times of her life as her doomed daughter's caretaker, never wavering in her displays of largely unreciprocated affection.⁹

Man's Best Friend: Kevin's seeing-eye dog Orson is a wonderful companion. But he has a streak of emotional volatility, and gets depressed when deprived of his nightly pampering. While it sometimes interferes with his other relationships, Kevin is completely dedicated to Orson, and never misses an opportunity to do right by his bud. Kevin's friends and

9 For discussions of Pedro Almodovar's greatest film, from which some features of the example are drawn, see Eaton, *Talk to Her*, and Shpall, "The Men of *Talk to Her*." For an insightful discussion of love for infant children, see Kennett, "True and Proper Selves."

family sometimes chastise him for what they see as a misguided loyalty, which often deprives them of his pleasant company. Kevin listens to their complaints respectfully but remains committed to his way of life. After all, he thinks, Orson is a joyous, innocent creature, whose affectionate help makes it much easier to navigate the world. And he fills Kevin's life with a unique sort of companionate satisfaction that personal relationships, with their inevitable strains and complications, have never really afforded him.¹⁰

The Book: After years of illness, and inner turmoil about the value of his dandified existence, Marcel decides to give up his engagement in the life of upper-crust Parisian society in order to devote himself to the composition of a massive literary work. Marcel is resolute in executing this decision. He sequesters himself at home and spends his remaining years pursuing his grand artistic project.¹¹ He has an ecstatic conception of the activity of writing and of his book's value. Indeed, Marcel sees the aesthetic achievements of literary composition as his only chance at a happy and meaningful life.¹²

The Way of Perfection: Teresa grows up in a pious household, and at an early age is captivated by stories about the lives of the saints. When her mother dies the young girl undergoes a transformative period of grief. Convinced that worldly concerns are futile, she takes up residence in a monastery. Teresa devotes her life to the perfection of divine worship,

- 10 "My dogs are my friends and part of my family. I know them better, actually, than I know the man living across the street. I do what I can to care for them and to keep us safe and well. They share more of my life than do my human friends" (Safina, *Beyond Words*, 287).
- 11 "So far from going into society, I would not even permit people to come and see me at home during my hours of work, for the duty of writing my book took precedence now over that of being polite or even kind. . . . I should have the courage to reply to those who came to see me or tried to get me to visit them that I had, for necessary business which required my immediate attention, an urgent, a supremely important appointment with myself" (Proust, *Remembrance of Things Past*, 3:1034–35).
- 12 "The idea of Time was of value to me for yet another reason: it was a spur, it told me that it was time to begin if I wished to attain to what I had sometimes perceived in the course of my life, in brief lightning-flashes . . . at those moments of perception which had made me think that life was worth living. How much more worth living did it appear to me now, now that I seemed to see that this life that we live in half-darkness can be illumined, this life that at every moment we distort can be restored to its true pristine shape, that a life, in short, can be realized within the confines of a book! How happy would he be, I thought, the man who had the power to write such a book! What a task awaited him!" (Proust, *Remembrance of Things Past*, 3:1088). See Nehamas, *Only a Promise of Happiness* and *On Friendship*, for some of the most wonderful reflections on the love of art in contemporary philosophy.

embodied in prayer. Eventually she gains a reputation for her wisdom and spiritual attainment, and, at the encouragement of church elders, develops theories about the methods and aims of ecstatic absorption with the deity.¹³

I hope that these examples depict the varied objects and life circumstances that occasion the experience of meaning. I hope also that they begin to motivate my conviction that this experience is most reliably caused by three main phenomena: life structure, susceptibility to rich and spontaneous emotion, and susceptibility to pleasing states of engagement and affection. This is the bedrock of the tripartite theory I will elaborate on in later sections. For now I offer some preliminary observations.

Max's love for Sherry structures his life. He devotes more time and energy to promoting Sherry's good than to addressing the flourishing of other beings and things he does not love; he spends more time with Sherry than with others; he cares more about getting her what she wants. Exactly the same can be said of Nancy and Kevin, though Alicia's sad fate will change this for Nancy.¹⁴ And while the cases of Marcel and Teresa may seem different, these appearances are superficial.

It is true that the activity of writing cannot feel pain. But it is nonetheless plausible that Marcel's writing can go better or worse, and it is plausible that Marcel's commitment to writing involves dedication to its progressive development. (Parallel claims about his book are also compelling.) Moreover, it is obvious that Marcel's craft structures his life, and that he may care more about writing than about even his closest friends and relations.

Some readers may look upon Teresa's case as a curiosity, because they believe that her love is in some sense undermined by an illusory belief in God's existence. But it is unfair to *assume* that Teresa's beloved does not exist. And it is worth observing that many of us love on the basis of profound illusions. Whether or not

13 "Now it seems to me that, when God has brought someone to a clear knowledge of the world, and of its nature, and of the fact that another world (*or, let us say, another kingdom*) exists, and that there is a great difference between the one and the other, the one being eternal and the other only a dream; and of what it is to love the Creator and what to love the creature . . . when one understands by sight and experience what can be gained by the one practice and lost by the other, and what the Creator is and what the creature, and many other things which the Lord teaches to those who are willing to devote themselves to being taught by Him in prayer, or whom His Majesty wishes to teach—then one loves very differently from those of us who have not advanced to that stage" (Teresa of Ávila, *The Complete Works of Saint Teresa of Jesus*, 27).

14 As I will note later, my account of love suggests interesting results about the possibility of loving the dead and other nonexistent objects, and about the nature of grief.

God exists, there is no doubt that commitment to Him structures Teresa's life as pervasively as any commitment structures most of ours.¹⁵

So the first important similarity among my cases is that love involves an intense commitment that crowds out commitment to other things. The examples of promoting the beloved's good and spending time with the beloved are two special but non-exhaustive cases of love's intense and constitutive form of commitment, which I call *devotion*, and explore in what follows.

A second important similarity is that the relevant psychological condition accounts for many of each of the protagonist's strongest emotions. It would be condescending, and clearly false, to insist that Kevin, Marcel, and Teresa could not have their most powerful emotional attachments directed at Orson, the craft of literary composition, and God, respectively. (Whether their attachments are rational or laudable is a separate question.) The emotional sensitivity that is rooted in intense devotion I call *vulnerability*.

Third, the devotion of my protagonists is a response to, and expressive of, an affectionate engagement with the object in question. In other words, they find pleasure or satisfaction in engaging with their beloveds. This observation is important for explaining how their cases differ from others that do not reliably lead to the experience of meaning. In my terminology, *liking* something is a key to meaning-generation.

Some readers may find my ideas so far unremarkable. I have not claimed that my views about the objects of love are wildly original. I have claimed that moral philosophers often ignore them. I will now defend this assertion more directly. Once I have shown that the notion of love that interests me is distinct from the notion of love that has interested most contemporary philosophers, I will proceed to explore it in more psychological detail.

3. LOVE FOR PERSONS

Contemporary philosophical inquiry into the nature and normative status of love has been dominated by a focus on the admittedly central case of love for persons. This focus has been both implicit and explicit. I have elsewhere given arguments for thinking the general tendency is worth interrogating.¹⁶ In this

15 Does God benefit from the love of believers? Now that is a chestnut. Either way, contemporary philosophy would benefit from sensitivity to cases like Teresa's, which have at other points in history been at the core of humanity's thinking about love, meaning, and morality.

16 Shpall, "Love's Objects." Compare also Plato's *Symposium* and Freud's homage (see the citation at Lear, *Love and Its Place in Nature*, 140): "Libido is an expression taken from the theory of the emotions. We call by that name the energy . . . of those instincts which have to do with all that may be comprised under the word 'love.' The nucleus of what we mean by love

section I hope mainly to establish two simple claims: first, that a narrow focus has been widespread, and second, that it has resulted in accounts of the nature of love that cannot apply to many of the cases I have taken to be paradigms of meaning-generation. Having done this, I will then criticize one idea that, if true, might threaten my approach: the idea that personal love is sufficiently normatively distinctive to make it the only kind of love truly relevant to moral philosophy.

Again, my primary intention is not to criticize other writers individually. They have various sensible reasons for offering more restricted accounts of love. The aim is to illustrate the relative advantages of my attempt to understand love via its connection to meaningfulness, without making stipulations about its possible or appropriate objects.

Here are a few representative examples of prominent philosophical theses about love.¹⁷

On J. David Velleman's Kantian conception, love is an arresting awareness of value in a person, a condition of intense emotional vulnerability that responds to another person's rational will. While Velleman notes that his characterization of love can also capture examples of vulnerability to other objects, he does not endeavor to say much about such cases.¹⁸ The distinctively Kantian idea presented in his luminous essay—that love is a moral emotion directed at instantiations of rational agency—could not be easily employed in such an explanation. In later work, Velleman emphasizes the distinctiveness of love for persons, and contrasts it with what we often call love for other objects, which he labels “benevolent affection.”¹⁹ Indeed, Velleman here argues explicitly that so-called love for dogs is rooted in an illusion: the illusion of personhood in our furry friends.²⁰

For Niko Kolodny, love is a particular mode of valuing, the valuing of a rela-

naturally consists (and this is what is commonly called love, and what the poets sing of) in sexual love with sexual union as its aim. But we do not separate from this—what in any case has a share in the name ‘love’—on the one hand, self-love, and on the other, love for parents and children, friendship and love for humanity in general, and also devotion to concrete objects and to abstract ideas. . . . We are of the opinion, then, that language has carried out an entirely justifiable piece of unification in creating the word ‘love’ with its numerous uses. . . . By coming to this decision, psycho-analysis has let loose a storm of indignation, as though it has been guilty of an act of outrageous innovation. Yet it has done nothing original in taking love in this ‘wider’ sense. In its origin, function, and relation to sexual love, the ‘Eros’ of the philosopher Plato coincides exactly with the love-force, the libido of psycho-analysis.”

17 Many more can be found in the helpful set of references in Helm, “Love,” which explicitly restricts discussion to personal love.

18 Velleman, “Love as a Moral Emotion,” 365.

19 Velleman, “Beyond Price.”

20 See Velleman, “Beyond Price,” 203, and, for criticism of this commitment, see Kennett,

tionship. Kolodny is clear that the relationship that concerns him, and that in his view occasions the special interest of moral philosophers, is one that necessarily involves persons:

On the one hand, I understand love exclusively as a state that involves caring about a person. However, it is perfectly correct English to say that someone “loves” something that is not a person. . . . “Love,” in ordinary usage, attaches to more than the psychological state with which I am concerned. . . . My narrowed focus is, I hope, acceptable. The species of love that involves caring for another person is the species that most attracts the interest of moral philosophers.²¹

Bennet Helm argues that love is intimate identification, or a kind of “taking to heart” of the beloved’s identity, where this identity is understood to be something like the perspective of an autonomous human agent.²² Troy Jollimore conceives of love as a special mode of seeing a person, which involves sympathetic and appreciative contact with the beloved’s qualities, and especially with her subjective experience of the world.²³ Though Helm and Jollimore recognize that we often take ourselves to love things that are not persons, they believe that this tendency is unfortunate. Love for persons is psychologically and normatively distinctive, and deserves its own philosophical theory.²⁴

More restricted targets are also common. Kate Abramson and Adam Leite explore the variety of love that they conceive of as a reactive emotion, which is a response to a good will.²⁵ Other philosophers give accounts of romantic love, a special kind of intimate love for an agent’s character, or, less commonly, close friendship.²⁶

None of these treatments can apply to the cases in which Kevin, Marcel, and Teresa figure. Nancy’s case is also troubling: Claire is not a person, and will never become one. It would be odd to think that Nancy’s love for her can be described in terms of appreciating rational agency or valuing an interpersonal relationship.

“True and Proper Selves,” 219–20. See also Millgram, “Kantian Crystallization,” for a general critique. I discuss Setiya’s similar treatment (in “Love and the Value of a Life”) below.

21 Kolodny, “Love as Valuing a Relationship,” 136–37.

22 Helm, *Love, Friendship, and the Self*.

23 Jollimore, *Love’s Vision*.

24 See Helm, *Love, Friendship, and the Self*, 2, and “Love”; and Jollimore, *Love’s Vision*, xii.

25 Abramson and Leite, “Love as a Reactive Emotion.”

26 On romantic love, see, e.g., Brogaard, *On Romantic Love*; de Sousa, *Love*; and Jenkins, *What Love Is*. On intimate love for an agent’s character, see Bagley, “Loving Someone in Particular.” On close friendship, see Nehamas, *On Friendship*.

These observations show that my focus is uncommon. Before offering my account of love, however, it may be worthwhile to reflect on the normative distinctiveness of personal love, since it is very likely judgments about its distinctive value that explain the shape of our discourse.²⁷

Love for persons is morally distinctive in at least the following ways. It tends to be more valuable than many or all other forms of love. Call this the *average value claim*. And it is in fact one of the most common kinds of love in human life. Call this the *prevalence claim*. These truths are important. Yet they do not cast doubt on the contention that other forms of love are also valuable and central sources of meaning. Indeed, these truths are compatible with the controversial idea that many people should pursue forms of nonpersonal love more actively, sometimes in place of personal love—an idea that I find plausible, significant, and underappreciated by philosophers and nonphilosophers alike.

What sort of judgment about value would explain and vindicate a philosophical discourse about love that ignored nonpersonal cases, if the average value and prevalence claims are insufficient? The relevant claim might be that love for a person necessarily realizes a distinctive kind of value—distinctive in the sense that the value of any other kind of love cannot be compared to it, or is necessarily inferior. These *distinctive value claims* are hard to formulate adequately and hard to evaluate. I am unsure which of my interlocutors accepts them.²⁸ Moreover, I do not need to deny them in order to convince the reader that my approach is worthwhile. Even if Marcel's love could not be compared to Max's love, or was necessarily inferior to it, I would remain convinced that Marcel's love is important and that moral philosophers should be interested in it. Nonetheless, I am inclined to reject distinctive value claims, and will say a few words on this score.

A first reason for doubt is that it can be appropriate to abandon personal love. This indicates that personal love may be less valuable than other things. Consider an abused woman who loves her abusive partner. Suppose she leaves him, finally resisting her impulse to stay after a long struggle with the recognition that her love is toxic. One description of her leaving: she has compared her love for her partner to other values that it threatens—for example, her love for her cats, her

27 See the Kolodny passage above; Velleman, "Beyond Price," where it is claimed that personal love is especially "fierce"; Helm, *Love, Friendship, and the Self*, where it is claimed that personal love is especially "deep"; and Bagley, "Loving Someone in Particular," 480, who claims that a particular form of personal love, grounded in appreciation of character, is the "best kind of love." Compare also Fromm, *The Art of Loving*, 20, for whom interpersonal love is "the mature answer to the problem of existence."

28 But compare Kant's view that rational agency is the sole ground of unconditional value (*Grounding for the Metaphysics of Morals*). For discussion of this connection, see Shpall, "Love's Objects."

love for her work, her love for herself—and judged her romantic love to be less important.²⁹ The aptness of this description would raise questions about both the comparison and inferiority versions of our idea about personal love's distinctive value.

Second, I am inclined to endorse the view that some people are unworthy of love. It was inappropriate for some who loved the adult Hitler to love him. It would have been far better for these Hitler lovers to love other things instead. The example is divisive and will not move everyone. For the sympathetic, however, it may serve as a reminder that some personal love is not just imprudent, but regrettable or even evil.

Third, I should emphasize the dialectical situation for proponents of a distinctive value claim. Their contention is that love for a person necessarily possesses value properties that no other kind of love can possibly possess. The negation of their claim is a far more modest possibility thesis.

Finally, I present a thought experiment for the reader to chew on, which indicates to me that many people already accept such possibility theses. For simplicity's sake, I limit myself to one comparison of personal love with love for aesthetic objects such as artworks. But further examples of similar kinds might be contemplated involving putative love objects such as infants, humans in profoundly deteriorated cognitive states, nonhuman animals, activities, works of art, deities, cities or nations, natural wonders, cultural endeavors, and social or political causes.³⁰

Imagine the following distant variant of a “one thought too many” case.³¹ In order to save the collection of paintings in the Van Gogh Museum, which is about to be set afire by an accomplished arsonist, an art-loving museum curator must immediately rush to intercept the perpetrator. But doing so will force her to miss a critical medical procedure that cannot be postponed. The procedure

29 Love for oneself is love for a person. But it is unclear which of the most prominent accounts of love can capture it. Compare Soble, “Concerning Self-Love,” and Frankfurt, *The Reasons of Love*.

30 You might think invoking the cases of love for infants or people in degenerated states in this particular argument is cheating, since philosophers interested in personal love must be interested in such cases as well. This is too quick. Most leading accounts cannot be extended so easily. For discussion, see Shpall, “Love's Objects.” For remarks on the love of country and of nature, see Lewis, *The Four Loves*, ch. 2. For more questionable examples, see Nelson, *Bluets*, on love for the color blue; and Charles, “A Crime of Passion,” on the love of yogurt. (Thanks to Michael Hardimon for the latter reference.) What distinguishes the boundaries of the lovable? I hope the tripartite theory gives us some purchase on this question without settling it too hastily. One set of ideas about the possibility of loving humanity will arise later, when I have said more about the nature of devotion.

31 Williams, “Persons, Character and Morality.”

would extend the curator's life for around six to twelve months; without it, she will die within the year. Choosing to save the paintings is, in other words, also choosing to shorten her life, choosing to spend less time with her husband and children and grandchildren, and so on. My conjecture is that at least some art lovers in this position would sacrifice their personal loves for the sake of the Van Goghs. (If you prefer a more mundane analogue, consider the claim that some artists would choose the completion of a beloved work over an extension of their own lives.)

Kieran Setiya makes a subtly different claim about moral distinctiveness, which might be thought more compelling than the idea I have just discussed.³² I conclude this section by quickly addressing it. According to Setiya, the bare property of being human justifies love, and specifically justifies love of a form that involves partiality in action.³³ It can be rational, for example, to act from love and save one's wife from drowning instead of saving three strangers. Now Setiya is surely correct that love involves, and to some extent justifies, partiality in action. But is this feature of justified partiality connected to love of human beings in particular, or love as such? It seems to me obvious that love for human beings is not distinctive in having implications for justified partiality. If Setiya is right about the permissibility of saving his wife instead of three strangers, it should likewise be permissible for him to save his puppy over three strange puppies, or the lone copy of his beloved unpublished manuscript over three strangers' similarly beloved and endangered works. The distinctive value here is general, inhering in love for many types of objects.³⁴

In sum: love for persons is one of the most valuable, common, and interesting forms of love. Human lives usually merit preservation and care more than other valuable objects do. But these facts do not make me any less interested in understanding love as a general psychological kind. On the contrary, they make me more curious about how to integrate love of different sorts into a flourishing life that is filled with the experience of meaning. Authors interested in the special case of love for persons are, as I have made clear, perfectly entitled to analyze it and it alone. We should nonetheless recognize that the dominance of

32 Setiya, "Love and the Value of a Life."

33 Though he is much indebted to Velleman ("Love as a Moral Emotion"), Setiya makes humanity rather than personhood the locus of his conception of love and its reasons, in order to address cases like *Talk to Her*.

34 Some think it is always impermissible to, for example, save one dog instead of one human being, independently of whether you love the dog. I am not sure. Consider Kevin in a lifeboat case with Orson and a terminally ill stranger. It is worth noting that this popular impermissibility claim would, if true, raise serious moral questions about pet adoption, which involves a huge allocation of resources to a pet, and which many of us take to be permissible.

this approach might limit us.³⁵ I have been at pains to motivate the availability of a wider understanding because the views I will now advance are in many ways responses to a discourse dominated by an exclusive focus on love for persons.

4. A TRIPARTITE THEORY

I now present a theory of love that systematizes the central features of my paradigm cases, and provides a richer characterization of the psychological phenomena that could play the meaning-generating role animating this inquiry. My strategy is to decompose love into three main elements: devotion, vulnerability, and liking. That these are imperfect terms of art should be obvious.³⁶

The three components are separable. They arise in different forms and degrees. Nonetheless, thinking about their individual natures and their potential combinations clarifies the differences between love and other psychological conditions.³⁷

The most important feature of my account is that I conceive of love as constituted by intense devotion. The devotion is of a particular character, which I will analyze shortly. Crucially, it is devotion that renders vulnerable and expresses liking.

- 35 For one quick illustration of this danger, consider the idea that love is essentially *a form of union*, which has been pursued in different ways by Scruton (*Sexual Desire*), Solomon (*About Love*), Nozick (“Love’s Bond”), and others (including its foundational presentation by Aristophanes in the *Symposium*). The idea may seem plausible when we only consider certain central cases of love for persons—and particularly the case of erotic love, which is the focus of most union theorists. (There are compelling worries about even this particular case. Healthy romantic love does not swallow up our interests or our autonomy. See Helm, “Love,” and Bagley, “Loving Someone in Particular.”) But it will be especially hard to vindicate a view like Scruton’s, which has it that there is no distinction between the lover’s interests and the interests of his beloved, when the beloved is an elephant or a national park. An expanded focus on a variety of examples, including cases of love for nonhuman animals and non-sentient objects, can be a valuable source of insight, and can inform our understanding of love for persons as well.
- 36 There are affinities here with the work of the psychologist Robert Sternberg, “The Triangular Theory of Love.” I cannot discuss the relationship between my account and Sternberg’s Triangular Theory at any length but will note that Sternberg is, like many contemporary philosophers, only interested in love for persons.
- 37 A reviewer worries that there is a tension between my composite characterization of love and my claim that love is a distinctive psychological condition. But various psychological conditions are composite yet distinctive. There are connections here to work in the psychology and philosophy of emotion, work that is too often neglected in writing about love. A useful introduction to some central debates is Griffiths, “Current Emotion Research in Philosophy.”

I hope that the virtues of this account will become apparent to the reader. But I will mention three at the outset: (1) It is an elaboration of an attractive moral psychology of love, one that is more detailed than many predecessors.³⁸ (2) It is controversial and distinct from existing views—even when it resembles them. (3) It gives us traction on puzzling intuitions about love’s connection to rational agency, as well as new insights about choice points for further theorizing.

One way to frame my picture of the nature of love, and to compare it to competing accounts, is by appealing to a divide between cognitive, affective, and conative psychological states. There are legitimate worries about this taxonomy of the mind, and I will not attempt to argue for its defensibility.³⁹ I will assume that most of us take belief, sadness, and desire to be somewhat helpful models of the cognitive, affective, and conative, and hope that readers will permit me to import these classifications without being overly encumbered by theoretical baggage. As should become clear, nothing much turns on this particular mode of division, though some of the most essential claims of alternative theories of love can be clarified by appealing to it.

For example, what Helm calls “Robust Concern Views,” like those of Frankfurt, conceive of love as a form of noninstrumental caring.⁴⁰ Proponents and critics alike understand such views as claiming that love is neither affective nor cognitive, but conative-volitional. Others think love involves cognition, in the form of value judgment—plausibly Velleman and Kolodny, but compare also Jollimore, who takes love to be an appreciative way of seeing a person.⁴¹ By contrast, the prevailing commonsense view is probably that love is a feeling or a mode of affect. Velleman’s analysis of love’s “arresting” of our emotional defenses is arguably a vision of this affective phenomenology, which Kant thought it so

38 Apart from some of the views already mentioned, compare Liao, *The Right to Be Loved*, ch. 4, and Ebels-Duggan, “Against Beneficence,” 143, for compositional accounts with less articulated content.

39 See Helm, *Emotional Reason*; Nussbaum, *Upheavals of Thought*, esp. ch. 1; and Fricker, *Epistemic Injustice*, 72–81.

40 Helm, “Love”; Frankfurt, “Autonomy, Necessity, and Love” and *The Reasons of Love*.

41 Velleman, “Love as a Moral Emotion”; Kolodny, “Love as Valuing a Relationship”; Jollimore, *Love’s Vision*. This assumes, against noncognitivists, that judgments of value express beliefs. Velleman (“Love as a Moral Emotion,” e.g., 360–61) is an especially fascinating example of a writer who endorses a cognitive requirement on love, since he regards love as an acknowledgement of the value of the beloved independent of motives to benefit and be with, and, as I note later, views conative analyses of love as “aim-inhibited versions of Freud” that “replace the sexual aim identified by Freud with the aims of desexualized charity and affection.” Velleman-love is a Murdochian exercise in “really looking,” and hence essentially involves cognition, perception, and belief.

important to reject—though Velleman’s view is complex in its blend of judgment and affect, and there are subtleties in trying to characterize it adequately.⁴²

My aim in introducing these distinctions between mental phenomena is minimal. The division I hope to exploit is the division between psychological states that involve our agency and those that do not—or, more precisely, between degrees of agency involvement. The key claim I will be invoking is the claim that devotion is a paradigmatically active or volitional condition. This distinguishes it from liking and vulnerability, which are much less active, even if they likewise involve or are associated with conation. It also distinguishes devotion from cognitive conditions like belief, awareness, and value judgment.⁴³

Devotion

Being devoted to something is being especially concerned about it, and being disposed to act on this concern. It is having the disposition to choose it over other things. Metaphorically, being devoted is having the beloved object near the center of your volitional universe. To pilfer a phrase from Iris Murdoch: the most robust loving is a life-occupation.⁴⁴

This is intentionally vague because it is common to implausibly narrow and moralize the nature of love’s devotion, which in fact takes a variety of forms. A popular thought is that love involves a tendency to promote the well-being of the beloved, which is rooted in noninstrumental concern for them.⁴⁵ And this is indeed one kind of devotion, an essential part of many recognizable forms of love, including nonpersonal love. For example, it is customary for art lovers to be devoted to art itself, and not simply to personal experiences of art. This is why

- 42 Velleman, “Love as a Moral Emotion.” Compare Kant, *The Metaphysics of Morals*, 450: “Love is not to be understood as feeling . . . or delight. . . . It must rather be thought as . . . active benevolence.” Velleman-love is, as I understand it, an arresting (affect) awareness (cognition) of the value of rational nature, as instantiated in a particular rational being.
- 43 Provided we reject cognitivist views about intention, which claim that intentions are special kinds of beliefs (e.g., Harman, “Practical Reasoning”; Velleman, *Practical Reflection*; Setiya, “Cognitivism about Instrumental Reason”).
- 44 “All this may sound ridiculous. But being in love is a life-occupation. I suppose this concept resembles, or rather is a special case of, the idea of doing everything for God and making the whole of life into a sacrament” (Murdoch, *The Black Prince*, 204). My characterization will accommodate Velleman’s point that we should be skeptical about positing necessary connections between loving and desiring *particular outcomes* (“Love as a Moral Emotion,” 361), while maintaining that love is most essentially a matter of the will.
- 45 See Taylor, “Love”; Frankfurt, *The Reasons of Love*; Rorty, “The Burdens of Love”; and many others. For an excellent discussion of how the noninstrumental concern of lovers relates to ancient Greek eudaemonism, see Brink, “Eudaimonism, Love and Friendship, and Political Community.”

many wealthy art lovers fund museums and scholarships instead of acquiring private collections. Similarly, loving pet owners like Kevin often act in ways that frustrate their own interests but further the good of their pets.

But there are other kinds of loving devotion. As Aristotle emphasized, close friendships are characterized by a commitment to shared activity.⁴⁶ As he did not emphasize, they are also characterized by devotion to the satisfaction of desires, even (at times) when such desire satisfaction detracts from flourishing.⁴⁷ My view is that three forms of devotion are central to love, and distinct from one another: devotion to flourishing, end-promotion, and being-with.⁴⁸ I think it is important to remain open-minded here. There is no fact of the matter about how much of each form of devotion is required for any particular form of love.

Plausibly only rational agents have their own ends, and this is one way in which appropriate love for them must be distinguished from appropriate love for other things. (Though I would stress that it is not obvious what beings count as rational agents, and also that many beings have desires whose satisfaction we might promote even if they do not have ends.) Then again, there are reasons for thinking that a certain species of being-with devotion is fundamental for meaning-generation. Though I cannot explore this proposal in detail, I think that Kiran Setiya is correct in arguing that the best response to one kind of midlife crisis is the adoption of “atelic” ends—or atelic orientations toward some of our activities.⁴⁹ This means conceiving of activities as completed or fulfilled at every point of our engagement with them. As Setiya nicely puts the contrast between the telic and the atelic: you might write a philosophy paper not in order to finish or publish it, but in order to be doing philosophy. My thought is that *being with our beloveds* is the gold standard of atelic activity, which is what makes it the right kind of response to crises of meaning.⁵⁰

46 Aristotle, *Nicomachean Ethics*.

47 See Nehamas, “The Good of Friendship,” who argues that philosophical accounts of friendship are excessively moralized, because friendship is often a matter of sharing in trivial, banal, or even reprehensible activities. Compare also Cocking and Kennett, “Friendship and Moral Danger.” And see Lewis, *The Four Loves*, 97: “Friendship (as the ancients saw) can be a school of virtue; but also (as they did not see) a school of vice. It is ambivalent.”

48 See Ebels-Duggan, “Against Beneficence,” for an interesting argument that love for fully functioning adults is a matter of devotion to their ends. The account is intended to outline a normative ideal, and so does not directly challenge the view that all three forms of devotion are partially constitutive of love, as I understand it. Bagley, “Loving Someone in Particular,” proposes an object of devotion in some loving relationships that he calls “deep improvisation,” and that may well be distinct from the three forms of devotion I explore.

49 Setiya, “The Midlife Crisis.”

50 Compare also Wonderly, “Love and Attachment,” on the value of self-interested attachment in romantic love.

Some of the most influential writers on love are skeptical about linking it to the will noncontingently. Velleman famously argues that conative conceptions are implausible descendants of a bunk Freudianism.⁵¹ Others agree, claiming that the active phenomena I have outlined are simply typical but non-necessary manifestations of love.⁵² Putative counterexamples to conative accounts include love for grown children and annoying relatives, which are said to sometimes lack any active orientation of will, such as a concern to help or promote well-being.⁵³ Perceptive critics have noted that these examples are under-described and dialectically weak.⁵⁴ Loving parents remain devoted to their adult children in all sorts of meaningful, if not grasping, ways. Our psychological connections to extended familial relations are extremely varied, and require careful differentiation before they can provide any general truths about love's nature. And, as I have just been arguing, love's devotion is not reducible to a moralized concern for the beloved's well-being. Even if they are not devoted to promoting other end-states (which I very much doubt), loving parents of self-sufficient adults are still devoted to spending time with their children.

Much more could be said about these examples and others like them. I cannot hope to establish the thesis that devotion is, as I believe, the core of love. But at a minimum my treatment suggests one way of conceptualizing a crucial choice we face in thinking about love's nature. Here is one way to put it. I can grant that there is some sense of love that is not active. But I maintain that my sense of love is the one that is most intimately connected to the perception of meaning. By way of illustration, let us consider a stronger case for the opponent of conative accounts—a case that deserves much more attention than it has received. To put it bluntly: Can we love the dead?

My claim is that the psychological condition most connected to the perception of meaning in life is essentially active. It follows that if it is possible to love the dead in my sense, then it must be possible to be devoted to them. I believe this to be a plausible consequence. If the dead have interests, then it is possible to be devoted to these interests. This may suffice for being devoted to the persons

51 Velleman, "Love as a Moral Emotion."

52 See Badhwar, "Love." Nussbaum agrees: "Love is a particular kind of awareness of an object, as tremendously wonderful and salient, and as deeply needed by the self. The project of possession (or of helping) is then a response to that awareness" (*Upheavals of Thought*, 477). Susan Wolf is instructively ambivalent on the issue. At times (e.g., *The Variety of Values*, 188) she expresses skepticism akin to Velleman's. But elsewhere (*Meaning in Life and Why It Matters*, 9–10, 26) she emphasizes that love must be "active" in order to adequately connect with meaningfulness.

53 Velleman, "Love as a Moral Emotion."

54 E.g., Helm, "Love"; Abramson and Leite, "Love as a Reactive Emotion."

themselves. (I am unsure.) Independently, some *believe* that such devotion is possible, because they believe in a personal afterlife. These people can presumably be lovingly devoted to those who have passed away. They are in Teresa's situation, if we suppose that God does not exist.

However, it is much harder to find meaning in a purely backward-looking, largely passive condition of remembrance. This does not imply that the life of a person without living loves is meaningless. There is of course some meaning to be found in remembrance, and in other psychological conditions, experiences, and relationships.⁵⁵ But the creation of new meaning in life is most characteristically the product of a special form of devotion, and the absence of devotion causes the deterioration of the experience of meaning. Some readers will find this point obvious. Those who want more evidence might reflect on familiar concerns about the psychological health of elderly people, particularly those separated by death from their friends and partners, and by distance from their living relatives.⁵⁶

These reflections cohere with another commonplace intuition. We are skeptical when someone claims to love something but has exhibited little devotion to it. Statements of the form, "If you really loved X, then you would Y," typically invoke actions or intentions that the agent in question apparently lacks, and are our default formulations of love-denial.⁵⁷ The force and ubiquity of these skeptical judgments is evidence that love is not merely contingently related to devotion. Again, this argument does not and could not *prove* that my conception of love is the best one available. No such arguments are forthcoming for any controversial view about this difficult topic. The best we can do is to gather up interesting evidence for rival conceptions, and evaluate this evidence in fresh ways.

It may be worth observing how my treatment of devotion distinguishes the idea of love articulated here from some influential traditions. Love in my sense involves sustaining an object near the center of your volitional universe, so it must be sharply distinguished from respect, goodwill, impartial benevolence, what I would call loving-kindness, and common understandings of *agape* that are connected to some of these notions.⁵⁸ Still, there might be some nonmeta-

55 I have nowhere maintained that *only* love can generate the experience of meaning.

56 Grief is in part a response to the loss of a focal point for one's volitional universe. It is no surprise that a strategy for coping is the adoption of new ends in the form of pets, activities, and cultural appreciation. The special susceptibility of the elderly to the loss of love and meaning, particularly in societies structured like ours, is a matter of great moral consequence.

57 See Fromm: "No assurance of [a mother's] love would strike us as sincere if we saw her lacking in care for the infant, if she neglected to feed it, to bathe it, to give it physical comfort. . . . It is not different even with the love for animals or flowers" (*The Art of Loving*, 25). Compare Ferracioli, "The State's Duty to Ensure Children Are Loved" 6.

58 See Nehamas, *On Friendship*, 50–51. Compare Aristotle: "Goodwill would seem to be a

phorical sense in which love for humanity (or love for elephants, or love for the cinema) can be expressed in devotion to, say, particular individuals, activities, and projects. This issue has divided philosophers for a long time, and I will not presume to settle it.

I conclude this discussion of devotion with a puzzle. The mental state of intention involves a practical commitment.⁵⁹ Indeed, it is natural to regard intention as the volitional state *par excellence*. So we might try to understand devotion as a robust pattern of intentions involving the beloved, which make up a holistic and directed structure of the will.⁶⁰ But this account may not suffice. Consider:

Untimely: Florentino loves Fermina madly. But he cannot devote himself to her in any ways that might be observed by others, since Fermina is married to Juvenal, and any displays of Florentino's love would have terrible consequences.⁶¹

The worry is that my emphasis on devotion requires me to deny that love can be hidden or unexpressed.

One important response is to insist that if Florentino truly loves Fermina, his intentions will in fact be structured around her, even in these sad circumstances. For example, Florentino may be devoted to thinking about Fermina every night, praying or planning for opportunities to meet her surreptitiously, and remaining romantically uncommitted.

But I find the spirit of the worry persuasive. So I prefer to analyze devotion as a combination of occurrent intentions and dispositions to intend. My formulations throughout this section have respected this ambiguity. Some of Floren-

feature of friendship, but still it is not friendship. For it arises even toward people we do not know, and without their noticing it, whereas friendship does not" (*Nicomachean Ethics*, 143). Since God has an infinite will, he may have mankind (or all Creation) at the center of His volitional universe. We must be more discriminating.

59 On Bratman's influential view (*Intention, Plans, and Practical Reason*, 107–10), intentions control conduct and structure deliberation. Intending to A now normally leads one to try to A; intending to A in the future normally survives to become an intention for the present. And an intention to A characteristically persists through time and exerts pressure against reconsidering whether to A. Additionally, an intention to A structures deliberation by disposing one to reason about means to A, and by disposing one to refrain from forming new intentions that are incompatible with A. See also Harman, *Change in View*, 94–95.

60 Compare Frankfurt, "On Caring" and *The Reasons of Love*, but with a more pluralistic account of the content of the relevant intentions, and the caveat that Frankfurt might reject my claims about vulnerability and liking.

61 "'Fermina,' he said, 'I have waited for this opportunity for more than half a century, to repeat to you once again my vow of eternal fidelity and everlasting love'" (García Márquez, *Love in the Time of Cholera*, 64).

tino's devotion may be masked by his unfortunate circumstances, just as the fragility of a glass may be masked when it is packaged. Still, these partially masked dispositions to intend to hang out with Fermina, promote her interests, and so on, are the heart of his love.⁶²

Now for our puzzle. I have just argued that devotion can be constrained or denied expression by external contingencies. It is hard to devote oneself in the normal ways while held in solitary confinement, for example.⁶³ Since most of us believe that confinement or separation does not always extinguish love, we are committed to the plausible view that, insofar as devotion is required for love, it can at times be realized in mere dispositions. However, there also appear to be limits to the adequacy of such purely dispositional claims. The puzzle concerns how we are to draw these boundaries.

Suppose a heroin addict feels that he loves his infant son, and has various dispositions to be devoted to him. Nevertheless, this man invariably ignores the child, overcome by, or giving in to, his stronger dispositions to get high.⁶⁴ I believe we should say, in at least some sufficiently bleak cases of this kind, that the man does not love his son, or does not love him in the sense of love articulated by the tripartite theory. I suspect that the root of this reaction is due to the swamping dispositions being internal to the agent—he is to some degree directly responsible for them—rather than being imposed from the outside.⁶⁵ But I am not confident in this idea, cannot defend it here, and will simply leave

62 Compare Goldie, "Love for a Reason," and Naar, "A Dispositional Theory of Love," on love as a disposition.

63 For extraordinary reflections on a more familiar example, see Aristotle, *Nicomachean Ethics*, 124–25: "Just as, in the case of the virtues, some people are called good in their state of character, others good in their activity, the same is true of friendship. For some people find enjoyment in each other by living together, and provide each other with good things. Others, however, are asleep or separated by distance, and so are not active in these ways, but are in the state that would result in the friendly activities; for distance does not dissolve the friendship without qualification, but only its activity. But if the absence is long, it also seems to cause the friendship to be forgotten; hence the saying, 'Lack of conversation has dissolved many a friendship.'" Soon after, Aristotle continues: "Those who welcome each other but do not live together would seem to have goodwill rather than friendship. For nothing is as proper to friends as living together." This marvelous passage has brought me an odd mix of painful recognition and consolation.

64 See Naar, "A Dispositional Theory of Love," 347, for discussion of a similar case involving depression.

65 Compare Aristotle, *Nicomachean Ethics*: "Voluntary action seems to be what has its principle in the agent himself" (32). Some readers have claimed that we cannot control dispositions, and thus worried that by countenancing dispositions to intend I give up on the volitional character of devotion. This cannot be right. Present-directed intentions are themselves dispositions. The relevant distinction is between dispositions to intend that we can

the reader to consider whether the embryonic distinction I have drawn has any significance. I hope it is clear that the puzzle is not necessarily an objection to my view. If the case is genuinely perplexing, then illuminating accounts of love will reveal why, rather than dissolving the puzzle without trouble.

I will make one final observation about Florentino's case. I agree that he lacks many of the forms of devotion normally constitutive of romantic love. But we should notice how much this explains. What Florentino wants is to be able to express his love for Fermina by devoting himself to her. Insofar as he cannot do this, or is substantially constrained in the ways he can do it, his situation is lamentable. After all, there is a special pain in unexpressed love, beyond a lack of reciprocation.⁶⁶ Even sure-to-be-unrequited love often seeks expression. The tripartite theory gives voice to this dynamic drama. If we lose all opportunities to "express our love," we eventually lose the love itself.⁶⁷

Vulnerability

Vulnerability is the least controversial element of my account. It is most commonly associated with Velleman's articulation of love's phenomenology, which outlines a special form of emotional receptivity that involves the breakdown of mechanisms of self-protection.⁶⁸ What I call vulnerability is implied by Velleman-vulnerability, but my claims about its nature are less ambitious.

To be vulnerable to an object of love is to be especially disposed to have strong emotions (for instance joy, heartbreak, pride, shame) conditional on the obtaining of states of affairs in which the object figures. When something bad happens to a beloved we feel pain, and we feel pain in part because we care about the beloved in her own right. Here we have an undeniable gulf between love and other relations of association. Those of us who are not saints might be pained by the terrible fates of casual acquaintances, but this resembles the parallel pains of love like shadows resemble forms.⁶⁹ To put the point in its more classical formu-

endorse and sustain, and those (e.g., unconscious desires) whose motivational operation does not involve our agency in these ways.

66 See Austen: "In vain have I struggled. It will not do. My feelings will not be repressed. You must allow me to tell you how ardently I admire and love you" (*Pride and Prejudice*, 128).

67 Compare another enigmatic Murdoch metaphor: "We cannot really love the dead. We love a fantasm that secretly consoles" (*The Black Prince*, 342). I already noted how we could avoid this conclusion, but it is worth stressing that my view makes it far from trivial to explain love for the dead in some cases, and that I think this is a virtue.

68 Velleman, "Love as a Moral Emotion" and "Beyond Price."

69 There is a sense in which we genuinely value strangers. However, since we remain largely unperturbed by many distant tragedies, it seems obvious that it is in another sense that we value what we love. Plausibly this is because "value" is ambiguous: sometimes it refers to

lation: we feel bound to what we love, and in an important way we share its fate.⁷⁰

To be vulnerable is to have lost control. Someone you love might move away, or start treating you with meanness, or shower you with new warmth, without your having done anything to precipitate it. Lovers are necessarily exposed to largely unmanageable sources of happiness, devastation, and the like. So while we may exercise substantial control in coming to love, loving itself renders us less emotionally autonomous. Jeanette Kennett puts it this way: “Love takes us hostage to fortune; it binds us to the weal and woe of the beloved in ways we could not have anticipated and cannot reject.”⁷¹

But vulnerability is not unique to personal relations. Losing a pet, suffering a career-ending injury, or confronting the demise of one’s long-term artistic or political ambitions can be devastating—more lasting and existential in its heart-break, for some, than losing a beloved grandmother or boyfriend.⁷² It may be irrational to have some of these patterns of vulnerability, but nobody said that human beings are never misguided in love. What seems clear is that love is partially constituted by these emotional dispositions. If you are not much moved by the fortune of something, this seems like pretty conclusive evidence that you do not love it.⁷³

The tripartite theory’s distinctive and controversial claim about vulnerability is that it is *explained by* devotion. I conclude this section with an argument for this explanatory claim. The argument is familiar from the philosophy of emo-

judging valuable, and other times it refers to caring. Only intense forms of caring render us vulnerable in the sense I am after.

70 The formulation recalls “union” views without endorsing them. Compare Prov. 14:10: “The heart knoweth his own bitterness, and a stranger doth not intermeddle with his joy.” We can interpret this wonderful sentence as expressing the view that only those who love us really feel for us. (Although I suspect this may be to interpret it erroneously, if we mean to widen the ambit of non-strangers beyond God!) A more direct and equally beautiful illustration, which I owe to Gabriel Citron, is a famous story about Rabbi Aryeh Levin: “And indeed, when his own good wife Hannah felt pains, he went with her to Dr. Nahum Kook and told him, ‘My wife’s foot is hurting us’” (Raz, *A Tzaddik in Our Time*, 150).

71 Kennett, “True and Proper Selves: Velleman on Love,” 217. See also Nehamas, *Only a Promise of Happiness*, 57, and *On Friendship*, 136. There are interesting comparisons to Frankfurt’s conception of volitional necessities (e.g., *The Importance of What We Care About*, ch. 7) that I cannot explore.

72 See Philippa Foot’s excellent discussion of “deep happiness” (*Natural Goodness*, 87), and her claim that even “things that do not really matter—like a bad *faux pas*, or the non-arrival of an invitation to a party of the Duchesse de Guermantes—can create any amount of disturbance, right up to obsession.”

73 Although here again we must worry about the types of masking cases I discussed in framing the puzzle about devotion.

tion, but it is not often invoked in writing about love.⁷⁴ It supports the general claim that emotions are concern-based. In other words, they are grounded in or explained by concern.

Why do I fear for my garden's fate in the coming storm? I am afraid because I value my garden's flourishing. Why am I over the moon about my sister's engagement? I am joyful because I care about my sister's happiness. Why am I upset about gerrymandering? I am upset because democratic ideals matter to me. Generalizing: concern (or something in the ontological neighborhood) is part of what explains emotion.

The argument looks good *prima facie*. Since vulnerability as I understand it is a disposition to have strong emotions, the concern-based nature of emotions would make it likely that vulnerability is also concern-based. I have maintained that devotion is an especially strong form of concern (practical commitment, intention, structure of will). It seems plausible, then, that intense vulnerability is explained by intense devotion.⁷⁵

Liking

Liking is the third element of the tripartite theory. That it is substantially less agential than devotion is the prevailing and intuitive view.⁷⁶ This contrast goes a long way toward explaining our complex and ambivalent intuitions about love's relationship with rational agency. I return to this point in my concluding remarks.

Many philosophers think that we can love without liking.⁷⁷ Concerning the sense of love under consideration, which is constitutively tied to meaning, I think they are wrong. Again, a virtue of the tripartite theory is that it clarifies this conceptual fault line. I hope that the ensuing reflections also provide some new considerations that bear on how we should proceed.

To like something is to be disposed to enjoy it, feel affection for it, experience attraction to it. There are deep connections between liking and desiring. If you like skiing, then, other things equal, you desire to ski.⁷⁸ It is a commonplace

74 Cf. Roberts, *Emotions*.

75 Compare Helm, *Love, Friendship, and the Self*, and Smuts, "Normative Reasons for Love, Parts I and II," 510.

76 Cf. Liao, "The Right of Children to Be Loved," 427.

77 See Velleman, "Love as a Moral Emotion"; Setiya "Love and the Value of a Life"; Wolf, *The Variety of Values*, 190; and Frankfurt, *The Reasons of Love*, 42. Cocking and Kennett ("Friendship and the Self," 519) and Nehamas (*On Friendship*, e.g., 109, 132) think that liking is part of loving, at least in the case of friendship. Badhwar, "Love," is an insightful proponent of the general love-liking connection. But explicit endorsements of the liking condition are rare, and arguments for it are almost never articulated.

78 Discussions in Mill, *Utilitarianism*, and Kenny, *Action, Emotion, and Will*, highlight similar

in philosophical psychology to treat desire as the fundamental pro-attitude, the state with world-to-mind direction of fit that we need, in addition to belief, in order to explain human action.⁷⁹ But liking and wanting are different, and I think it is preferable to concentrate on the relationship of liking to loving.⁸⁰

Liking something is compatible with finding it frustrating and with a host of other negative emotions. The claim here is not the manifestly absurd one that love involves unqualified hedonic stimulation. The claim is that devotion, even devotion that renders vulnerable, only plays the robust meaning-generating role if you like the object to which you are devoted.

In ordinary English, loving and liking appear intimately connected. But the surface of our language might be confusing. One source of skepticism about the merits of a pluralist account of love's objects is the perceived cheapness of some love attributions. The word love plausibly has multiple functions, some of which fail to cut the mind at its joints. For instance, it commonly picks out relatively superficial states ("I love this mint chip ice cream") that seem more like bare enjoyment, and unhelpful in thinking about the love that interests us here.⁸¹

Still, our linguistic intuitions may provide us with some information. The claim that you love a nonhuman object but do not like it has a paradoxical ring. If you love writing fiction then you like writing fiction—though you might not like particular aspects of it, and some stretches of writerly life may be hard going.⁸² If I am right that liking is not merely incidental to loving in such cases, then this is a piece of evidence for any analysis of love that aspires to object-generality.

Suppose Max's brother Mark claims to love Sherry, but it appears that he does not like spending time with her, does not take any satisfaction from the maintenance of their relationship, and has mostly negative reactions to even thinking about her. There is something unsettling but also incongruous about this apparent combination of attitudes. And it seems importantly different from

connections between pleasure and desire.

79 See, e.g., Smith, *The Moral Problem*, and Sinhababu, "The Humean Theory of Motivation Reformulated and Defended." Frost ("On the Very Idea of Direction of Fit") has recently attacked the philosophical convention of explicating mental states partially in terms of direction of fit. I cannot engage with his stimulating arguments here, and the reader need not accept the convention in order to make sense of my views.

80 Berridge, "Wanting and Liking."

81 For this worry, see Helm, *Love, Friendship, and the Self*, 2, and "Love." Rorty, "The Burdens of Love," 347, seems too quick to take our language as dispositive.

82 Fiction writers sometimes claim that their craft is pure anguish. This seems hyperbolic insofar as they also claim to love writing. Machado describes personal love in a way that I think generalizes: "We love each other all of the time and like each other most of the time" ("Mothers," 55).

Max's psychological orientation in *World's Best Mom*, which I offered as a paradigm case of filial love. Though we do sometimes make claims about loving persons without liking them, it is not clear that we should take these assertions at face value. For example, "I don't like you, but I love you" may just convey something like, "I love you, but you've been acting like an asshole."⁸³

Nonetheless, many dismiss the love-liking connection. In the remainder of this section I consider a case that I take to best express their worries. The discussion will permit me to reflect on some of my dissatisfactions with common arguments, and will hopefully illustrate some distinctive features of the tripartite theory.

Lost Cause: Jill has always been a model mother to her son Beelzebub. Sadly, Beelzebub's moral deterioration has reached a point of no return. Jill remains devoted to her son. She does everything she can to help him and to spur some kind of change in his perspective and way of life. However, she is consistently unsuccessful in bringing about meaningful change, and is tortured by the situation. After years of soul-searching, Jill finally confesses to her analyst that even the last dregs of motherly affection have disappeared.⁸⁴

Some philosophers are explicitly committed to claiming that Jill loves Beelzebub, even though she does not like him. And it is plausible that many philosophers who do not discuss examples of this sort nonetheless implicitly endorse the conclusion, since they conceive of love as a specific kind of emotional vulnerability, attachment, union, identification, care, or valuing, and each of these conditions can in principle be directed at objects we do not like.

83 This is how I interpret the iconic phrase of Smokey Robinson's in "You've Really Got a Hold on Me." (Thanks to Gary Watson for the fantastic reference.) A related example, which I owe to Jacob Ross, is from David Bowie's "Drive-in Saturday": "She's uncertain if she likes him/But she knows she really loves him." Poetry is often forged from twists in our conceptual expectations. We can see something interesting in these lines without thinking they constitute an objection.

84 See Gaitskill: "I love her—I love her dearly—because I'm her mother and I can't help it. But I don't like her" ("Heaven," 202). Compare Velleman: "It is easy enough to love someone whom one cannot stand to be with" ("Love as a Moral Emotion," 361). Another common and related case is that of young siblings: we sometimes say that they love each other, even though they "cannot stand one another." Here is my brief interpretation of this assertion by way of an analogy. I really like *Vertigo*. Nonetheless, if you were to play this film on my television every morning for a dozen years, I might at some point begin to say that I cannot stand it. In other words, overexposure to things that we like makes us inclined to say that we dislike them; and in some real ways we do. But it is very unclear what to conclude. *Lost Cause* is a less noisy case that better expresses the essence of the disagreement.

My claim is that liking the objects of our devotion is a critical part of finding our interactions with them to be reliable sources of meaning in life. I cannot defend this view except by employing imaginative exercises. So I ask the reader to imagine that her life is filled with devotion that renders vulnerable but does not express liking. In order to do this, she might begin by imaginatively inhabiting Jill's position, and generalizing it across the space of her intimate relations.⁸⁵ She might then attempt to inhabit the unfortunately all-too-common experience of vulnerable devotion to a career that she does not like. And so on. Writ large upon the canvas of a life, this psychological condition seems to me a case study in existential malaise.

I have one more argument for the liking component, which is error-theoretic, i.e., a diagnosis of why we may be inclined to mistakenly reject it. The diagnosis is that we think of Jill's vulnerable devotion as especially praiseworthy. Our commitment to its moral worth infects our judgments about whether Jill loves Beelzebub—which is unsurprising, given that “love” often functions as a kind of moral honorific. Refusing to admit Jill into the class of lovers can seem cold. It feels like a refusal to bestow moral approval where it is due. This natural thought should be resisted, though. There are many praiseworthy orientations of mind that moral psychology should distinguish.

The availability of this error theory supports my skepticism about largely unsupported intuitions that sever loving from liking. And we should recall that cases like Jill's are relatively unusual, and hard to describe and interpret. Whatever we think about whether Jill should count as a lover in some sense, I take myself to have provided reasons for thinking that she is not a lover in the sense of interest here, in which love is the condition most robustly connected to the experience of meaning.

5. CONCLUDING REMARKS

In presenting the tripartite theory, I suggested that it could explain and perhaps

85 For a grave comparison, consider Andrew Solomon's profile of Peter Lanza, the father of the Sandy Hook killer. Here is how the piece (“The Reckoning”) concludes:

I wondered how Peter would feel if he could see his son again. “Quite honestly, I think that I wouldn't recognize the person I saw,” he said. “All I could picture is there'd be nothing there, there'd be nothing. Almost like, ‘Who are you, stranger?’” Peter declared that he wished Adam had never been born, that there could be no remembering who he was outside of who he became. “That didn't come right away. That's not a natural thing, when you're thinking about your kid. But, God, there's no question. There can only be one conclusion, when you finally get there. That's fairly recent, too, but that's totally where I am.”

dissolve puzzles about the connections between love and rational agency. These puzzles concern perennial topics of interest in philosophy, literature, and the arts. Is love an emotion? Is it under our control? Are we responsible for it? Is it governed by reason, and subject to evaluation? Are its demands antagonistic or complementary to the demands of morality?

These questions are enduring objects of fascination partly because we have never agreed about what love is. In providing a detailed account of love as a psychological kind, the tripartite theory furnishes us with new materials for answering them. Even opponents of the theory may find it useful for identifying the at times murky fault lines dividing opposing conceptualizations of this elusive psychological condition.

To conclude, I offer a pregnant observation about one of these themes, which has been anticipated at various points above.

Puzzling intuitions about control and responsibility animate thinking about love in everyday as well as philosophical discourse. Compare the following representatively incompatible passages from Roger Scruton and Robert Solomon:

Erotic love, like the love of children, is compelled by the embodiment of its object. . . . We are subjected by erotic love. . . . Our freedom suffers the impact of an external necessity. Erotic love is experienced, not as a decision, but as a destiny.⁸⁶

Love is a *decision*. A decision *to* love, and a decision about *whom* to love, and how, and when, and why. Romantic love is an emotion of *choice*.⁸⁷

We commonly presuppose something like Scruton's view that love—especially erotic love—is akin to a compulsion, or even a sickness that overwhelms our agency.⁸⁸ And this leads us to make ambitious assumptions about responsibility and justification, for example concerning the inaptness of blame.⁸⁹ Yet we also

86 Scruton, *Sexual Desire*, 233.

87 Solomon, *Love*, 212.

88 Compare Cervantes: "Perhaps you consider me a man whose power of reasoning is weak and, even worse, one who has no judgment at all. It would not be surprising if that were the case, because it is evident to me that in my imagination the power of my afflictions is so intense and contributes so much to my ruination that I am powerless to prevent it and I become like a stone" (*Don Quixote*, 234). And see how Wilcox, "Love's Coming," expresses the presupposition: "She had looked for his coming as warriors come,/with the clash of arms and the bugle's call;/but he came instead with a stealthy tread,/which she did not hear at all."

89 See Lewis: "When lovers say of some act that we might blame, 'Love made us do it,' notice the tone. A man saying, 'I did it because I was frightened,' or 'I did it because I was angry,' speaks quite differently. He is putting forward an excuse for what he feels to require ex-

find ourselves deeply uncomfortable with these intuitions. Upon reflection, all does not seem to be fair, permitted, or excusable even in “matters of the heart.”⁹⁰

The tripartite theory allows us to make a simple conjecture about these deep perplexities. Our intuitions are muddled because love is a composite psychological condition, whose component parts are very differently susceptible to control, and very differently amenable to normative assessment.

Even the most ambitiously rationalistic philosopher will acknowledge that our fundamental likes and affinities are hard to manipulate directly. Something similar might be said about the intentional manipulation of our vulnerabilities, once we have come to have them. Whereas we may choose to devote ourselves to a child, a cat, or a project in a way that resembles how we may choose to raise an arm. These claims about control shape our intuitions about responsibility for love, and assessment of it, in predictably interesting ways.

Much more needs to be said to vindicate the importance of this final set of observations. I hope that this essay has paved the way by presenting a plausible and distinctive conception of love that illuminates to some degree the connection between love and meaning.⁹¹

The University of Sydney
sam.shpall@sydney.edu.au

cusing. But the lovers are seldom doing quite that. Notice how tremulously, almost how devoutly, they say the word *love*, not so much pleading an ‘extenuating circumstance’ as appealing to an authority” (*The Four Loves*, 136).

- 90 Compare Korsgaard: “I am using the term ‘grounds’ here not to avoid but rather to emphasize the obscurity of the *because* in love, which seems to fall somewhere in between the *because* of practical reason and the *because* of causality. The grounds of love do seem to have something in common with practical reasons . . . they are unlike mere causes, and like practical reasons, in that they can (sometimes) be right or wrong, or at least better and worse. . . . Yet the grounds of love do not quite seem to be practical reasons, and indeed seem to operate more like causes. To the extent that love is a passion, we do not decide to love on the basis of its grounds, for we do not *decide* to love at all. . . . For all of these reasons, the notion we are dealing with is an obscure one, in need of more philosophical attention” (“The General Point of View,” 8).
- 91 For especially helpful comments and encouragement, I would like to thank Zed Adams, Richard Arneson, Julia Borcherding, David Brink, Gabriel Citron, Stephen Darwall, Kory DeClark, Kenny Easwaran, Nate Gadd, Daniel Greco, Daniel Harris, Bennett Helm, Agnieszka Jaworska, Yao Lin, Dustin Locke, Errol Lord, Rachel McKinney, Eliot Michaelson, Shyam Nair, Aaron Norby, Alejandro Pérez-Carballo, David Plunkett, Sara Protasi, Daniel Putnam, Jacob Ross, Samuel Rickless, Jeffrey Seidman, Mark Schroeder, Gary Watson, Caroline West, Gideon Yaffe, Yuan Yuan, and an anonymous reviewer for *JESP*.

REFERENCES

- Abramson, Kate, and Adam Leite. "Love as a Reactive Emotion." *The Philosophical Quarterly* 61, no. 245 (October 2011): 673–99.
- Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett Publishing, 1999.
- Austen, Jane. *Pride and Prejudice*. 1813. Mineola, NY: Dover Publications, 1995.
- Badhwar, Neera. "Love." In *The Oxford Handbook of Practical Ethics*, edited by Hugh LaFollette, 42–69. Oxford: Oxford University Press, 2005.
- Bagley, Benjamin. "Loving Someone in Particular." *Ethics* 125, no. 2 (January 2015): 477–507.
- Berridge, Kent C. "Wanting and Liking: Observations from the Neuroscience and Psychology Laboratory." *Inquiry* 52, no. 4 (2009): 378–98.
- Bratman, Michael. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.
- Brink, David. "Eudaimonism, Love and Friendship, and Political Community." *Social Philosophy and Policy* 16, no. 1 (Winter 1999): 252–89.
- Brogaard, Berit. *On Romantic Love: Simple Truths about a Complex Emotion*. Oxford: Oxford University Press, 2015.
- Cervantes, Miguel de. *Don Quixote*. 1615. Translated by Edith Grossman. New York: Harper Collins, 2003.
- Charles, Dan. "A Crime of Passion: When the Love of Yogurt Burned Too Bright." National Public Radio, July 13, 2015.
- Cocking, Dean, and Jeanette Kennett. "Friendship and Moral Danger." *Journal of Philosophy* 97, no. 5 (May 2000): 278–96.
- . "Friendship and the Self." *Ethics* 108, no. 3 (April 1998): 502–27.
- de Sousa, Ronald. *Love: A Very Short Introduction*. Oxford: Oxford University Press, 2015.
- Eaton, A. W., ed. *Talk to Her: Philosophers on Film*. Abingdon, UK: Routledge, 2008.
- Ebels-Duggan, Kyla. "Against Beneficence: A Normative Account of Love." *Ethics* 119, no. 1 (October 2008): 142–70.
- Ferracioli, Luara. "The State's Duty to Ensure Children Are Loved." *Journal of Ethics and Social Philosophy* 8, no. 2 (September 2014): 1–19.
- Foot, Philippa. *Natural Goodness*. Oxford: Oxford University Press, 2001.
- Frankfurt, Harry. "Autonomy, Necessity, and Love." In *Necessity, Volition, and Love*, 129–41. Cambridge: Cambridge University Press, 1999.
- . *The Importance of What We Care About*. Cambridge: Cambridge University Press, 1988.

- . “On Caring.” In *Necessity, Volition, and Love*, 155–80. Cambridge: Cambridge University Press, 1999.
- . *The Reasons of Love*. Princeton: Princeton University Press, 2004.
- Fricker, Miranda. *Epistemic Injustice*. New York: Oxford University Press, 2007.
- Fromm, Erich. *The Art of Loving*. London: Unwin Books, 1971.
- Frost, Kim. “On the Very Idea of Direction of Fit.” *Philosophical Review* 123, no. 4 (2014): 429–84.
- Gaitskill, Mary. “Heaven.” In *Bad Behavior: Stories*, 177. New York: Simon and Schuster, 1988.
- García Márquez, Gabriel. *Love in the Time of Cholera*. Translated by Edith Grossman. New York: Knopf, 1987.
- Goldie, Peter. “Love for a Reason.” *Emotion Review* 2, no. 1 (2010): 61–67.
- Griffiths, Paul. “Current Emotion Research in Philosophy.” *Emotion Review* 5, no. 2 (2013): 215–22.
- Harman, Gilbert. *Change in View*. Cambridge, MA: MIT Press, 1986.
- . “Practical Reasoning.” In *Review of Metaphysics*, edited by Alfred R. Mele, 431–63. Oxford: Oxford University Press, 1976.
- Helm, Bennett. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press, 2000.
- . “Love.” *Stanford Encyclopedia of Philosophy* (Fall 2017). <https://plato.stanford.edu/archives/fall2017/entries/love>.
- . *Love, Friendship, and the Self*. Oxford: Oxford University Press, 2010.
- Jenkins, Carrie. “Modal Monogamy.” *Ergo* 2, no. 8 (2015).
- . *What Love Is: And What It Could Be*. New York: Basic Books, 2017.
- Jollimore, Troy. *Love’s Vision*. Princeton: Princeton University Press, 2011.
- Kant, Immanuel. *Grounding for the Metaphysics of Morals*. 1785. Translated by James W. Ellington. Indianapolis: Hackett Publishing Company, 1993.
- . *The Metaphysics of Morals*. 1797. Edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- Kennett, Jeanette. “True and Proper Selves: Velleman on Love.” *Ethics* 118, no. 2 (2008): 213–27.
- Kenny, Anthony. *Action, Emotion, and Will*. Abingdon, UK: Routledge, 1963.
- Kolodny, Niko. “Love as Valuing a Relationship.” *Philosophical Review* 112, no. 2 (2003): 135–89.
- Korsgaard, Christine. “The General Point of View: Love and Moral Approval in Hume’s Ethics.” *Hume Studies* 25, nos. 1 and 2 (1999): 3–41.
- Langton, Rae. “Duty and Desolation.” *Philosophy* 67, no. 262 (1992): 481–505.
- Lear, Jonathan. *Love and Its Place in Nature: A Philosophical Interpretation of Freudian Psychoanalysis*. New Haven, CT: Yale University Press, 1990.

- Lewis, C. S. *The Four Loves*. 1960. New York: HarperCollins, 2012.
- Liao, Matthew. "The Right of Children to Be Loved." *Journal of Political Philosophy* 14, no. 4 (December 2006): 420–40.
- . *The Right to Be Loved*. Oxford: Oxford University Press, 2015.
- Machado, Carmen Maria. "Mothers." In *Her Body and Other Parties*, 45–64. Minneapolis: Graywolf Press, 2017.
- Mill, John Stuart. *Autobiography of John Stuart Mill*. 1873. New York: Columbia University Press, 1969.
- . *Utilitarianism*. 1861. Edited by George Sher. Indianapolis: Hackett Publishing, 2001.
- Millgram, Elijah. "Kantian Crystallization." *Ethics* 114, no. 3 (April 2004): 511–13.
- Murdoch, Iris. *The Black Prince*. London: Penguin Classics, 2003.
- Naar, Hichem. "A Dispositional Theory of Love." *Pacific Philosophical Quarterly* 94, no. 3 (September 2013): 342–57.
- Nehamas, Alexander. "The Good of Friendship." *Proceedings of the Aristotelian Society* 110 (2010): 267–94.
- . *On Friendship*. New York: Basic Books, 2016.
- . *Only a Promise of Happiness*. Princeton: Princeton University Press, 2007.
- Nelson, Maggie. *Bluets*. Seattle: Wave Books, 2009.
- Nozick, Robert. "Love's Bond." In *The Examined Life: Philosophical Meditations*, 68–86. New York: Simon and Schuster, 1989.
- Nussbaum, Martha. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press, 2001.
- Shpall, Sam. "Love's Objects." In *Love, Reason, and Morality*, edited by Katrien Schaubroeck and Esther Engels Kroeker, 57–74. New York: Routledge, 2017.
- . "The Men of Talk to Her." *Film and Philosophy* 17 (2013): 96–112.
- Plato. *Symposium*. In *Plato: Complete Works*, edited by John M. Cooper, 457–505. Indianapolis: Hackett Publishing, 1997.
- Proust, Marcel. *Remembrance of Things Past*. 1913–1927. Translated by C. K. Scott Moncrieff, Terence Kilmartin, and Andreas Mayer. 3 vols. New York: Random House, 1981.
- Raz, Simcha. *A Tzaddik in Our Time: The Life of Rabbi Aryeh Levin*. Translated by Charles Wengrov. Jerusalem: Feldheim Publishers, 1976.
- Roberts, Robert C. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press, 2003.
- Rorty, Amelie. "The Burdens of Love." *Journal of Ethics* 20, no. 4 (December 2016): 341–54.

- Safina, Carl. *Beyond Words: What Animals Think and Feel*. New York: Henry Holt and Company, 2015.
- Scruton, Roger. *Sexual Desire: A Philosophical Investigation*. London: Weidenfeld and Nicolson, 1986.
- Setiya, Kieran. "Cognitivism about Instrumental Reason." *Ethics* 117, no. 4 (July 2007): 649–73.
- . "Love and the Value of a Life." *Philosophical Review* 123, no. 3 (July 2014): 251–80.
- . "The Midlife Crisis." *Philosophers' Imprint* 14, no. 31 (November 2014): 1–18.
- Singer, Irving. "From *The Nature of Love*." In *The Philosophy of Erotic Love*, edited by Robert C. Solomon and Kathleen M. Higgins, 259–78. Lawrence: Kansas University Press, 1991.
- Sinhababu, Neil. "The Humean Theory of Motivation Reformulated and Defended." *Philosophical Review* 118, no. 4 (October 2009): 465–500.
- Smith, Michael. *The Moral Problem*. Hoboken, NJ: Blackwell Publishing, 1994.
- Smuts, Aaron. "Normative Reasons for Love, Parts I and II." *Philosophy Compass* 9, no. 8 (August 2014): 507–26.
- Soble, Alan. "Concerning Self-Love: Analytic Problems in Frankfurt's Account of Love." *Essays in Philosophy* 12, no. 1 (January 2011): 55–67.
- . *The Philosophy of Sex and Love*. 2nd ed. St. Paul, MN: Paragon House, 2008.
- Solomon, Andrew. "The Reckoning." *New Yorker*. March 17, 2014. <https://www.newyorker.com/magazine/2014/03/17/the-reckoning>.
- Solomon, Robert. *About Love: Reinventing Romance for Our Times*. New York: Simon and Schuster, 1988.
- . *Love: Emotion, Myth, and Metaphor*. Amherst, NY: Prometheus Books, 1990.
- Sternberg, Robert. "The Triangular Theory of Love." *Psychological Review* 93, no. 2 (1986): 119–35.
- Taylor, Gabriele. "Love." *Proceedings of the Aristotelian Society* 76 (1976): 147–64.
- Teresa of Ávila. *The Complete Works of Saint Teresa of Jesus*. Vol. 2. Translated and edited by E. Allison Peers. London: Sheed and Ward, 1975.
- Velleman, J. David. "Beyond Price." *Ethics* 118, no. 2 (January 2008): 191–212.
- . "Love as a Moral Emotion." *Ethics* 109, no. 2 (January 1999): 338–74.
- . *Practical Reflection*. Princeton: Princeton University Press, 1989.
- Wilcox, Ella Wheeler. "Love's Coming." In *Poetical Works of Ella Wheeler Wilcox*. Edinburgh: W. P. Nimmo, Hay, and Mitchell, 1917.

- Williams, Bernard. "Persons, Character and Morality." In *Moral Luck*, 1–19. Cambridge: Cambridge University Press, 1981.
- Wolf, Susan. *Meaning in Life and Why It Matters*. Princeton: Princeton University Press, 2010.
- . *The Variety of Values: Essays on Morality, Meaning, and Love*. Oxford: Oxford University Press, 2015.
- Wonderly, Monique. "Love and Attachment." *American Philosophical Quarterly* 54, no. 3 (July 2017): 235–50.

PRACTICAL REASON NOT AS SUCH

Kenneth Walden

CONSTRUCTIVISTS THINK that value is a “construction” of the attitude of valuing. For a thing to be valuable, “is for that thing’s value to be entailed from within the point of view of a creature who is already valuing things.”¹ I think this is a compelling conception of value. But it is also a bleary view. For it is consistent with many different notions about *which* value claims are correct, *how* they are to be constructed from an agent’s practical point of view, and *what* constitutes such a point of view.

One front for these questions is the debate between Kantian and Humean constructivists. These philosophers disagree about whether we can discern anything substantive about the construction of value by looking at the activity of practical reasoning “as such.” A Kantian like Christine Korsgaard thinks we can. Using practical reason commits us to the Categorical Imperative, and from this we can derive specific duties and prohibitions. These normative claims can therefore be constructed from every practical point of view.² In contrast, a Humean constructivist like Sharon Street denies that practical reason as such commits reasoners to much of anything—that no “substantive moral conclusions are entailed from within the standpoint of normative judgment as such.” Instead, “the substantive content of a given agent’s reasons,” Street says, “is a function of his or her particular, contingently given, evaluative starting points.”³ So if you do not give a damn about pain on Tuesday or the suffering of others, and you can coherently and self-consciously maintain that indifference, then there is no ground for saying that you nonetheless have a reason to avoid that pain or ease that suffering.

This dispute acknowledges two potential grounds for normative correctness

- 1 Street, “What Is Constructivism in Ethics and Metaethics?” 367.
- 2 Korsgaard, *Sources of Normativity*. Another version of the view is defended by Markovits, *Moral Reason*, 145–62.
- 3 Street, “What Is Constructivism in Ethics and Metaethics?” 370. Defenses of views that could be reasonably called Humean include Street, “Constructivism about Reasons,” and Lenman, “Humean Constructivism in Moral Theory.”

within the set of factors that shape a person's behavior. First there is "practical reason as such." This is the universal and perfectly generic faculty for practical deliberation *tout court* and so a factor in nearly every action performed by a rational agent. (The notion seems to be a descendant of Kant's "pure" practical reason: our capacity for reasoning practically abstracted from any empirical and so contingent conditions.) And then there are what Street calls our "particular, contingently given" evaluative attitudes—my abhorrence of Torquemada's cruelty, my admiration for Pollini's phrasing, my preference for Darjeeling in the afternoon.

But surely there are elements of our moral psychology that do not fit neatly into these pigeonholes—things that are neither particular evaluative attitudes nor features of practical reason *as such*. For brevity's sake I call these "further factors." I worry that neglect of these features' role in shaping what is entailed from our practical point of view can keep constructivism from being as sophisticated and catholic a view as it might be.

This worry prompts the questions I take up here: Might there be constructivist arguments in support of distinctive normative judgments whose starting point is these further factors? And might such arguments constitute a constructivist program that complements those advocated by Street or Korsgaard? Bigger questions are at stake as well, ones I am sneaking up on by entering into this intramural dispute between constructivists. They are questions about practical reason *not* as such. Are there contingent features of agents that are nonetheless best understood as genuine facets of their faculty of practical reasoning? And, if there are such features, might they play a role in determining which normative judgments are correct for an agent in something like the way that practical reason "as such" is supposed to? The debate between different flavors of constructivism is a good stalking horse for these questions because constructivists agree on a tight connection between practical reason and normativity.

The goal of this article is to defend an affirmative answer to all of these questions. The first step in doing this is identifying what I call "further factors" and saying why we should believe in them. This I do in section 1. The second step is showing that they have some significance for which normative judgments are correct. I do this within the general setting of constructivism in section 2. In section 3 I suggest that this version of the view can contribute to the cause of establishing the extensional adequacy of constructivism. In section 4 I close with a brief discussion of the import of the notion of practical reason *not* as such.

1. FURTHER FACTORS IN THE PRODUCTION OF ACTION

I said that there are features of our moral psychology that contribute to the actions we perform but fall in between the two poles that Street's and Korsgaard's dispute centers on—between practical reason as such and our particular evaluative attitudes. In this section I present two examples.

My first example is an item from the social scientist's toolkit. Someone insults Igor, and Igor challenges him to a duel. Why does Igor do this? One part of our explanation cites Igor's evaluative attitudes and beliefs. He has some end, and he believes that challenging his antagonist is a means to that end. But suppose Igor's challenge is part of a larger pattern in Igor's community: a constellation of interconnected practices and symbols centering around honor that are inculcated early in life and have a pervasive influence on the way of life is carried out by those in that community. In this case it would seem a full explanation of Igor's action must say something more. It must say something about this regularity and Igor's role within it.

Now suppose that when Igor arrives at the museum he patiently waits in line to purchase a ticket. Why does Igor do this? Here, too, one part of the explanation cites Igor's evaluative attitudes and beliefs. But this also seems incomplete. For one, Igor's behavior instantiates a distinctive regularity. If he is like most of us, his queuing is not the result of any weighing of ends and calculation about how to achieve them. Rather, his initial recognition of the row of idle people is as a queue—as a *thing-to-be-waited-in*. As such he never seriously entertains the possibility of striding past it; this possibility is “silenced” in his deliberations.⁴ The practical problem Igor faces is framed in a particular way: *not* how to get into the museum most quickly, but how to enter the queue. There are exigencies that might spur Igor to skip the line, of course, but these exceptions prove the rule. Skipping the line would be “cutting,” i.e., a violation of the norm.

The obvious way to supplement our initial explanation of Igor's challenge is to say that Igor lives in an honor culture, that he has internalized a particular role in that culture, and that his challenge is prescribed by that role. Likewise, the obvious way to supplement our explanation of Igor's waiting in line is by saying that he has internalized a social norm, the norm of queuing, and that he is following that norm. When giving this sort of explanation the thing we are attributing to Igor is a *psychological schema*. As Sally Haslanger explains:

A schema consists in clusters of culturally shared concepts, beliefs, and

4 For a defense of the claim that social norms silence in this way, see Hlobil, “Social Norms and Unthinkable Options.”

other attitudes that enable us to interpret and organize information and coordinate action, thought, and affect. Schemas are public—think of them as social meanings conventionally associated with things in our social world, including language—but are also internalized and guide behavior.⁵

Our two examples of schemata differ along a few dimensions. The schema that Igor has internalized insofar as he is a “man of honor” is diffuse but pervasive: it informs many aspects of his life, but does so in relatively subtle ways. The queuing norm he has internalized, by contrast, is relatively localized in its application but highly specific in its guidance. They are, nonetheless, examples of the same basic phenomenon: the psychological internalization of some aspect of a social structure by an agent that directs her behavior in the ways necessary to conform to and reproduce that structure.

The action-guiding power of a schema is unique in a few respects. First, a significant part of its influence on our action consists not in directing our behavior in particular directions, as is characteristic of aims, ends, and projects, but by structuring what Haslanger calls a “choice architecture.” Schemata “structure the possibility space for agency” by silencing options, making others salient, and “providing templates of interaction that favor (or discourage) certain forms coordination with respect to a resource, e.g., share, hoard, distribute; and by canalizing our attitudes accordingly.”⁶ For example, an honor schema may silence overly conciliatory or artful ways of acting while making forceful and candid options more salient. It may offer templates for interaction between men and women whose employment allows both to maintain their status as honorable. And it may codify relatively well-structured rituals like the duel. This way of controlling behavior allows schemata to, as Pierre Bourdieu says, “generate and organize practices and representations that can be objectively adapted to their outcomes without presupposing a conscious aiming at ends.”⁷

Second, schemata direct behavior, in part, by penetrating an agent’s perceptual and cognitive systems. As the anthropologist Paul Friedrich says about honor in particular, “[it] is a code for both interpretation and action; in other words with both cognitive and pragmatic components. . . . Honor consists of a system of symbols, values, and definitions in terms of which phenomena are conceptualized and interpreted.”⁸ Thus, people who have internalized honor schemata

5 Haslanger, “What Is a (Social) Structural Explanation?” 126.

6 Haslanger, “What Is a (Social) Structural Explanation?” 128.

7 Bourdieu, *The Logic of Practice*, 53.

8 Friedrich, “Sanity and the Myth of Honor,” 284–85.

come to see particular acts as cowardly, attend to sleights of etiquette or deference, and instinctually associate masculinity with virtue in ways not typical of those lacking that schema.

Third, schemata are relatively resistant to change and updates—more so than, e.g., an agent's desires or aims. Changing or dislodging a psychological schema generally requires not merely *changing one's mind* about some judgment of fact or value, but *retraining* oneself in the modes of thought that the schema directs. As a result, there are notable examples of social norms and their attendant schemata persisting despite active and conscious resistance.⁹

Fourth, the guidance offered by schemata cannot be understood individually. That is, we cannot understand the way that the internalization of the queuing norm guides Igor's behavior without understanding that what he has internalized is something essentially *public*, as an instance of a norm that has been internalized by multiple agents for the coordination of their behavior. If Igor is the only person who queues (or ever has), then he is not acting out an internalized norm, but exercising a private caprice. Furthermore, schemata generally depend on publicly available material resources. The practice of queuing is a way to coordinate access to a scarce resource—a bank teller's attention, a ride on Space Mountain—and in turn uses the spatial and temporal resources that facilitate individual instances of queuing. This trade-off of resources must be equilibrated for the practice to be sustained. The explanatory use of psychological schemata brings these elements in train. If we say that Igor issued a challenge because he has internalized an honor schema, we are not *merely* locating a cause of his action within his own psychology (as we might understand an explanation in terms of beliefs and desires), but orienting Igor within a shared social practice.

If we want to cite schemata like these as examples of "further factors" in action, we face two questions. First, why should we countenance such things in the first place? The answer to this is that doing so gives us the best, most complete explanations of human behavior. In particular, relying on schemata offers a kind of explanatory power not found elsewhere. As Haslanger puts it, schemata "offer insight into why the particular individual behaved as he/she did, but [they] also contribute to our understanding of the individual as the instance of a type—a type defined by the conditions for existing at that node. By carving the explanandum across a broader range of possibilities (as a type, not a token), we can achieve better, more stable, explanations."¹⁰

The second question is harder. How do schemata stand relative to the contrast described above? Can they be reduced to some combination of evaluative

9 See Bicchieri and Fukui, "The Great Illusion."

10 Haslanger, "What Is a (Social) Structural Explanation?" 128.

attitudes and the activity of practical reasoning as such? Or are they an example of what I call further factors?

The first reductive possibility is easily rejected: schemata are not features of practical reasoning as such since they are clearly contingent factors in action. At first blush it seems equally obvious that they cannot be assimilated to evaluative attitudes either. Our paradigms of valuing are attitudes such as liking, appreciating, admiring, preferring, wanting, hating, fearing, and scorning. It is hard to see how the psychological work of, say, an honor schema could be reduced to one or even a set of these.

Now, it is true that schemata often *involve* evaluative attitudes: valuing one's honor is arguably a constitutive component of internalizing the honor schema, and one may be motivated to internalize a norm in order to avoid censure or to gain the advantages of coordination that norms provide. And it is also true that some valuing attitudes can display some of the features of schemata I listed a moment ago. Love may be capable of structuring our choice architecture by making certain actions "volitional necessities." Desires can be "backgrounded." Attitudes whose objects are social, like patriotism, may be unintelligible outside of a social context.¹¹ Nonetheless, it is hard to imagine fully reducing a schema to any array of desires, lovings, admirings, or fearings. Igor's internalization of the queuing norm is not just a matter of his valuing the coordination afforded by such norms, disvaluing the censure that awaits violators, valuing his self-conception as a rule-follower, or even some combination of these. We can imagine someone who has internalized the norm despite lacking these evaluative attitudes and someone who has failed to do so despite having them all. And even if we *could* produce a set of paradigmatically evaluative attitudes that collectively captured all of Igor's behavior that we associate with the schema, it is questionable whether this set would offer the same explanatory advantages as the schema itself. We may be able to identify the efficient cause of a particular action with this set, but it is not clear how it would afford us the structural understanding—understanding Igor's act as a token of a larger regularity—that citing the schema does.

Upon initial inspection, then, psychological schemata look like *bona fide* examples of further factors. This initial impression may be too hasty, though. Street emphasizes that she has a technical notion of valuing in mind that is more liberal and structurally articulated than the standard conceptions. She says this while explaining that valuing is a very different attitude from mere desire. Much of the resistance to attitude-dependence theories of value, she thinks, can be traced

11 See Frankfurt, "Autonomy, Necessity, and Love"; Pettit and Smith, "Backgrounding Desire"; and MacIntyre, *Is Patriotism a Virtue?*

back to the unfortunate assumption that relevant dependence basis is desire. This leads Street to recommend the attitude of valuing as a basis for her constructivism, and, more importantly, to suggest ways in which her notion of valuing is more inclusive and structurally complex than the ordinary notion of desire. It is possible that the assimilation of schemata to evaluative attitudes will look more promising once we understand this conception.

Street distinguishes valuing from mere desire in a few ways, but only one of them is relevant to our question. The attitude of valuing, she says,

is characterized by greater structural complexity than the attitude of mere desiring. We tend to think of “desiring” as directed at a single object or state of affairs: I desire a donut, for example, or to be rich or to be liked. Evaluative experience of the kind that confers value if anything does, however, is structurally a great deal more complicated than that. It often involves experiencing very specific features of the world as “calling for” or “demanding” or “counting in favor of” other very specific things. For example, I experience the fact that a friend lent me her car two months ago as counting in favor of saying “yes” to the favor she’s asking me now; I experience someone’s youth and inexperience as ruling out a harsh reply; and so on. Such states of mind are very different from simply wanting a donut. . . . The attitude of valuing involves much more complex attitudes toward the world and one’s own potential responses to it.¹²

This clarification is crucial, since a schema could simply *be* one of the “much more complex attitudes” that valuing issues in. Trouble is, Street does not say *what* these attitudes are, only what they are not necessarily. This makes it possible to read Street’s characterization of the attitude of “valuing” as infinitely capacious—as potentially encompassing nearly *every* factor in an agent’s behavior except the demands of practical reason as such. If we did this, then schemata would count as evaluative attitudes, but only trivially.¹³

Of course, how Street intends to use the word “value” is only of secondary importance. The real question concerns the “joints” of moral psychology: the theoretically important differences among the factors contributing to an agent’s

12 Street, “Coming to Terms with Contingency,” 43–44.

13 Bernard Williams’s characterization of the “subjective motivational set” is similarly open-ended: “It can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent” (“Internal and External Reasons,” 105). The items on this list are so heterogeneous as to make me think that Williams means to include any item that is “subjective” and has the potential to explain behavior. In that case I do not see any reason to think these items comprise a kind.

actions. Street is concerned with one such difference, between necessary and contingent factors. But this may not be the only or even the most significant division. So we should ask: Is there a theoretically significant difference in the way that our paradigms of evaluative attitudes—liking, admiring, fearing—influence an agent’s behavior in the way that psychological schemata do?

I think there is, and the best way to describe it is through an analogy suggested by Cristina Bicchieri:

Like a collection of linguistic rules that are implicit in a language and define it, social norms are implicit in the operations of a society and make it what it is. Like a grammar, a system of norms specifies what is acceptable and what is not in a social group. And analogously to a grammar, a system of norms is not the product of human design and planning.¹⁴

Let us flesh out this analogy a little bit by thinking about factors at work in a person’s linguistic behavior. First, there are the speaker’s communicative aims: asserting that water is clear, offering a model of canine speciation, promising to help you move. These aims are highly contingent and up to the speakers to adopt. Then there are the conditions of language *as such*: publicity, recursivity, etc. In between these two poles lie such things as the grammars of individual languages. That English has a particular grammar is a contingent thing, but it is contingent a very different way from Igor’s aim of using language to ask Mrs. Igor for a cup of tea. Likewise, speakers of a language have some control over their grammar—the grammar is what it is because of linguistic practices enacted by linguistic actors, and grammar can change over time—but this control is deeply attenuated when compared to the control they have over their own linguistic aims.

The internalization of a grammar shapes linguistic behavior in much the same way that psychological schemata shape behavior more generally. Grammar guides our linguistic behavior by structuring our “choice architecture”: not by directing us to perform particular speech acts, but by giving us templates for formatting those speech acts. It penetrates our perceptual and cognitive systems: I need not consult a rule to recognize the problem with subject/verb disagreement—it just *sounds* wrong. It is sticky: I cannot simply decide to change the grammatical rules that I follow, even if they can slowly evolve over time. And it is essentially social: it cannot be understood independently of its connection to a public language spoken by other people with whom I want to communicate.

I go to the trouble of describing these parallels because I think it is obvious that there is a significant theoretical difference—a joint—in how the internalization of a grammar contributes to an agent’s linguistic behavior and how

14 Bicchieri, *The Grammar of Society*, xi.

that agent's contingent linguistic aims do the same. (Even though they are both, strictly speaking, contingent features of language use.) And so it would be a mistake to try to assimilate the former to the latter. There is an equally important and structurally analogous difference between psychological schemata and our paradigmatic evaluative attitudes. The distinctively *grammar-like* way that schemata guide agents' behavior sets them apart from our paradigms of valuing, and this, I suggest, makes them good examples of further factors.

My second example of a further factor involves an agent's physical characteristics. Igor wants a sip of tea, so he stretches out his arm, grasps his cup, retracts his elbow, and imbibes. Why does Igor do all these things? Because that is the way to get the tea, and it is the way to get the tea because Igor cannot lift the cup with his mind or magically teleport the liquid from cup to gullet. Call the complex of features that explain our actions in this way our *embodiment*.

We face the same two questions with this example as we did with our first. Why should we believe in such a quality? And is it a genuine example of a "further factor"? The answer to the first question is much the same as before. Explanations that adduce the fact that Igor is a creature with arms and fingers oriented in space without telekinetic powers afford us a structural understanding of Igor's actions insofar as they allow us to understand those actions as part of a greater regularity that encompasses other, similarly embodied agents. This explanatory benefit is a reason to believe in the quality. As for the second question, there should also be little temptation toward reducing embodiment to a set of evaluative attitudes. And since we can well imagine different sorts of creatures who are capable of practical reason but have very different physical constraints on their agency, we should not think that it reflects the demands of practical reason as such.

Instead, the main source of resistance to the suggestion that Igor's embodiment represents an interesting further factor in his acting is that it is not really a feature *of Igor (qua agent)* at all, but a nonnormative fact that Igor must reckon with while deliberating. Igor sipped his tea on Main Street instead of Elm Street because that is where the café is. So this geographical fact is one factor in Igor's behavior. But this is not a feature *of Igor* the way that his evaluative attitudes and capacity for practical reasoning are. So why think differently about his embodiment? The suggestion implicit in this question is to assimilate Igor's embodiment to the undistinguished swathe of nonnormative facts that he encounters in his deliberations. His having a body and all the capacities and restrictions that come along with that have the same significance for his behavior as geographical facts about which street has a café. And so they are not "further factors" in any interesting sense.

Like the suggestion of assimilating schemata to the attitude of valuing, I think this proposal elides an important difference. We can get a handle on it by examining the phenomenology of choice. Suppose Igor is waiting for a flight and decides to have a cup of tea. To do this, he briefly considers all the shops near his gate, excludes those that do not sell tea, and heads off to the nearest one that does. By contrast, no one decides to have a sip of tea, entertains all logically possible ways of accomplishing that, and only *then* crosses off possibilities like teleportation that are incompatible with her embodiment. For agents embodied like us, the question of how to drink tea *just is* the question of how to bring one's body into the right kind of contact with tea, and not a special instance of a more general question that happens to apply to us.

One could reply that there are other potential explanations for this difference: the difference is explained by Igor's habituation to his body and that he could, in principle, become habituated to other things just as thoroughly. But the relationship between Igor's practical reasoning and his embodiment seems special in a few respects. For creatures like us, practical reasoning is in the *first instance* an embodied activity: a project that involves moving our bodies around in space and time. Our conception of practical reason *as such*—of a faculty for reflection and deliberation divorced from our embodiment—is an abstraction from this activity. This is a point Charles Taylor has emphasized. Our practical perspective on the world, he observes, “is essentially that of an embodied agent, engaged with or at grips with the world”—which is to say that “our perception *as an experience* is such that it could only be that of an embodied agent engaged with the world.”¹⁵ For example, the world we face when we set about to do things is “oriented vertically, [where] some things are ‘up,’ others are ‘down’; and in depth some are ‘near,’ others ‘far.’ Some objects ‘lie to hand,’ others are ‘out of reach’: some constitute ‘unsurmountable obstacles’ to movement, others are ‘easily displaced.’”¹⁶ And when we face this world, we do so under certain constraints. We face it as “an agent who acts to maintain equilibrium upright, who can deal with things close up immediately, and has to move to get to things farther away, who can grasp certain kinds of things easily and others not, can remove certain obstacles and not others, can move to make a scene more conspicuous; and so on.”¹⁷ Finally, our training in practical reasoning involves training in how to manipulate our bodies. The cultivation of a person's ability to recognize the abstract categories of means and ends begins with the infant's recognition that certain basic actions relieve certain forms of distress—that suckling relieves

15 Taylor, “The Validity of Transcendental Arguments,” 23, emphasis added.

16 Taylor, “Lichtung or Lebensform,” 62.

17 Taylor, “Lichtung or Lebensform,” 62–63.

hunger—and the capacity for intertemporal forms of reasoning, like planning, is rooted in the child’s ability to tolerate the visceral discomfort that accompanies delayed gratification.

These features distinguish an agent’s embodiment from the nonnormative facts she must deal with in her deliberations. We do not introduce embodiment as a special constraint on an antecedently abstract activity of practical reasoning, as we might introduce geographical facts about tea as constraints on instrumental reasoning toward the end of tea drinking. On the contrary, the more generic activity is secondary: it is an abstraction away from our usual way of going about practical reasoning, which is structured by our embodiment. This constitutes an important difference—a joint—between facts like a given café being on Main Street and those features of an agent that constitute her embodiment. In an important sense, then, an agent’s embodiment is not exogenous to her practical reasoning: it conditions that reasoning, rather than acting on it as an external force or constraint. And for this reason I think we should accept that it is a genuine example of a “further factor.”

In this section I have been trying to identify “further factors” in our action—things that contribute to our behavior but cannot be reduced to the items that Humean and Kantian constructivists have focused on in their projects. I have offered two examples: the way a social structure is psychologically internalized by an agent through a schema and the way an agent is embodied. Despite the hypothetical resistance I have considered here, I do not think the existence of these factors should be very surprising. The hard part of my case comes next: establishing that these things make a difference to what reasons an agent has.

2. THE NORMATIVE SIGNIFICANCE OF FURTHER FACTORS

To make this argument, I work within the broadly constructivist framework I adumbrated in the introduction. The primal idea of metaethical constructivism is that the correctness of normative facts is a function of what can be constructed from an agent’s practical point of view. This means that two things figure in normative correctness: the agent’s practical point of view—the construction basis—and the methods employed in performing the construction. For constructivists, then, the question of whether the further factors I identified in the previous section have normative significance is the question of whether they affect either of these components.

A *prima facie* case can be made for including further factors in either of them. Psychological schemata and forms of embodiment shape the way that agents view the world when engaged in practical thinking, and they do so in very deep

and persistent ways, so it seems fair to say they are part of the agent's "practical point of view." On the other hand, these factors structure agents' practical reasoning in distinctive ways: they structure their choice architecture, organize their experience into discrete practical problems, and make particular solutions to those problems salient. So insofar as our construction procedure is supposed to reflect the work undertaken by practical reason, it seems likely that further factors will play a role here as well.

There are some reasons to be skeptical about both possibilities, however. First, one could object that, for the purposes of normative construction, an agent's "practical point of view" ought to include only those attitudes that reflect her full-fledged normative judgments—her evaluative attitudes. Further factors may condition an agent's practical outlook without earning her endorsement, but precisely because they lack such endorsement we ought not see them as part of the practical point of view from which normative correctness is constructed. Second, in determining normative correctness we want to know what is *really* entailed by these attitudes, not what agents might *suppose* is entailed by them given their idiosyncratic styles of reasoning. This means taking practical reason as such as our standard, not the parochial procedures employed by imperfect agents.

Those moved by these objections are likely to favor a more austere constructivism. Street's program is a good example. For her, the practical point of view is "the set of all of the relevant agent's normative judgments, minus the normative judgment whose correctness is in question."¹⁸ Normative correctness is constructed from this point of view by a particular conception of the demands of practical reason as such: by standards grounded in the constitutive nature of the attitude of valuing as such.

Just as it is constitutive of being a parent that one have a child, so it is constitutive of taking oneself to have conclusive reason to *Y* that one also, when attending to the matter in full awareness, takes oneself to have reason to take what one recognizes to be the necessary means to *Y*. One cannot take oneself to have conclusive reason to *Y* without taking oneself to have reason to take the means to *Y*.¹⁹

These constitutive features lay down standards by which normative judgments can be deemed correct or mistaken. Thus, from the point of view of someone who takes himself to have a reason to *Z*, the judgment "I have a reason to *Y*," where *Y* is the necessary means to *Z*, is correct, and the judgment "I have no rea-

18 Street, "Constructivism about Reasons," 223–26. Here I follow Street's usage and treat "normative judgment" and "evaluative attitude" as equivalent.

19 Street, "Constructivism about Reasons," 228.

son to Y " is incorrect—just by dint of the constitutive standards implicit in the attitude of taking oneself to have a reason. Through the repeated application of these standards to an agent's evaluative attitudes, the entailments of that agent's practical point of view are constructed.

My argument for the normative significance of further factors is that more austere constructivists face certain problems precisely because they deny this significance. The *prima facie* case is easy enough to sketch, but it has to be filled out with examples. The standards that Street says are introduced by the constitutive nature of valuing include things like: if A values e , and m is a necessary means to e , then A values m ; if A values r , and $r = s$, then A values s ; if A values all x that are F and c is F , then A values c .²⁰ These are essentially the laws of first-order logic applied to the valuing operator plus one distinctively practical constraint, the instrumental principle. They are *very* undemanding standards, and this should make us worry that combining them with an agent's evaluative attitudes will not suffice to settle many of the questions about normative correctness that ought to be answerable.

Some examples will help bring this out. Suppose that Igor has internalized an honor schema and someone has insulted him. Does Igor have a reason to challenge this person to a duel? Intuitively, he does—or rather, that is what we expect someone who thinks that normative correctness is constructed out of an agent's practical point of view to say. But can we get this conclusion out of Street's constructivism? We may be tempted by the following argument:

1. Igor values his status as an honorable man.
2. Therefore Igor has a reason to maintain this status.
3. Issuing a challenge is the necessary means to preserve his status.
4. Therefore Igor has a reason to issue a challenge.

Here the conclusion is meant to follow from the constitutive standards of valuing in precisely the way Street describes. But there is a hitch: it is not clear that (3) is true. (Or even that Igor judges it to be true, though for Street what matters is the fact, not the agent's opinion.) For it may be possible for Igor to preserve his status by offering a trenchant diagnosis of what is wrong with the practice of dueling and convincing members of his community that honor is really about law and respect rather than vengeance. Or for Igor to preserve his status by finding some fiendishly clever way to humiliate the person who insulted him. Or for Igor to avoid losing face by failing to issue the challenge, but then doing something so stupendous that his honor is preserved all the same. These may be remote possibilities, but they serve to underscore the larger point. The condition on the

20 Street, "Constructivism about Reasons," 227–31.

standard that Street says is constitutive of valuing—being a necessary means—is quite strong, and that makes the normative force of the standard weaker than we might have expected.²¹

This is an example of the underdetermination I think Street's view faces. There are a few ways she might reply. The first is to accept the underdetermination while denying that it is a defect. Street is open to this possibility in some cases. Sometimes, she says, "the standards legislated by a person's other normative judgments, coupled with all the relevant non-normative facts about necessary means, etc., are insufficient to yield a result one way or another."²² This is a sensible suggestion when a person is ambivalent or apathetic—when his evaluative attitudes produce a "tie" between different judgments or he simply lacks the relevant attitudes—but it seems inapt here. Igor is neither ambivalent nor apathetic in this situation. Rather, the apparent underdetermination seems to arise from the constitutive standards of Igor's evaluative attitudes and the nonnormative facts failing to mesh in a way that settles the case. Moreover, we should be reluctant to bite the bullet and accept underdetermination in this sort of case, since it has the potential to generalize the innumerable similar cases in which an agent's behavior is mediated by an internalized schema.

The second response would have us try to overcome the underdetermination with further evaluative attitudes. That is, we could attribute enough evaluative attitudes to Igor such that, when combined with a sufficiently fine-grained picture of the nonnormative facts, the question of Igor's reasons would, after all, be settled by even a very weak conception of scrutiny. For example, we could observe that Igor's alternatives to issuing a challenge may require him doing things he is reluctant to try or lead to outcomes that he disfavors. Of course, we would also have to assess the likelihood of success in each instance. If we amend premise (3) in this way, it would look like an expected utility calculation: for all ϕ such that ϕ is a logically possible alternative to issuing a challenge, the probability that

21 There is a complication here that I am bracketing. Street defines normative correctness in an unusual way. Instead of giving a constructive definition of normative correctness, she gives a constructive definition of normative *mistakenness* and defines correctness as, essentially, non-mistakenness. Thus for her a judgment is, so to speak, correct until proven mistaken. (We can see how unusual this is by comparing it to mathematical constructivism, in which the analogous view would have it that a mathematical proposition is correct unless a counterexample can be constructed.) One consequence of this definition is that the underdetermination I am alleging shows up in a unique way. Instead of entailing that there are p for which neither p nor $\sim p$ is correct, it entails that there are p for which *both* p and $\sim p$ are correct since *neither* judgment violates the weak standards set out by the constitutive requirements of the agent's evaluative judgments.

22 Street, "Constructivism about Reasons," 236.

ϕ succeeds in preserving Igor's honor times the degree that he values ϕ -ing and its consequences is less than the probability that issuing a challenge preserves his honor times the degree to which he values that option and its consequences. Or something along those lines.

Rejiggering premise (3) like this could render the argument sound, but it would mean attributing a multitude of relatively nuanced evaluative attitudes to Igor—ones about all the various ϕ -ings and their potential consequences.²³ There are a few reasons not to take this path. First, it is not obvious why we are entitled to attribute such a wide range of subtly variegated evaluative attitudes to Igor. It is unlikely that he has consciously formed all of these attitudes, since many concern esoteric options. Rather, we probably feel justified in attributing them in a purely dispositional fashion. We need to attribute these attitudes, we might say, in order to explain Igor's behavior in a case exactly like the one we are thinking about. But this is not true. As we have seen, we can get a *better* explanation of Igor's behavior by saying that he has internalized a psychological schema associated with honor culture, and we can do so without attributing a bounty of evaluative attitudes. This explanation is better both because it gives us the structural insight into Igor's behavior that social explanations offer and because it is more psychologically parsimonious.

Second, this would mean understanding Igor's evaluative attitudes in a radically dispositional way—as consisting in anything we need to attribute to him in order to explain his behavior on a roughly Humean model of explanation. This seems to undercut the idea that was cited to motivate austere constructivism—namely that the practical point of view should be understood to include only those commitments genuinely endorsed, rather than factors that affect an agent's action without this kind of acceptance.

Finally, the proposal has the result that the construction of Igor's reason looks hairy, since it depends on something like a complex, expected value calculation, even though, for Igor, the existence of a reason to issue a challenge could not be more obvious. Of course, we should not insist that our construction procedure mimic an agent's actual reasoning, but in this case there is not anything to suggest that Igor's immediate recognition, without adverting to value calculations, represents an error in his reasoning.

For these reasons, I think we should be reluctant about overcoming the apparent indeterminacy this case reveals simply by adding as many additional evaluative attitudes as needed to produce a version of premise (3) that renders the argument sound.

23 Cf. Street, "Constructivism about Reasons," 233n42, in which she defends liberality about the attribution of evaluative attitudes.

The final possibility is to suppose that there are additional constraints on the project of scrutiny—the construction of normative correctness—than Street imagines. These can be grounded in a more robust conception of practical reason as such, or in some further, contingent feature of Igor. The former is the sort of proposal we might associate with Kantian constructivists: practical reason as such involves more than Street’s quasi-logical constraints, and these can help settle whether Igor has a reason to issue a challenge. In principle this strategy could work, but the actual proposals from Kantian constructivists seem poorly suited to the task. Korsgaard suggests that everyone is committed to valuing humanity as such. It is not clear how this particular value is going to help in Igor’s case, except, perhaps, by saying that the whole business of dueling is irrational. Markovits says that reason requires us to achieve a higher degree of coherence than logical consistency—what Kant calls “systematic unity” in the first *Critique*. I am not sure how this would help either, since the problem does not seem to be that Igor’s values are inadequately systematic, but that it is not clear what is *entailed* by a set of more or less systematic values.

That leaves one option for overcoming the underdetermination, that a contingent feature of Igor’s moral psychology, beyond his evaluative attitudes, bridges the gap. Here we find an obvious proposal: our original argument that Igor has a reason to issue a challenge is sound—and so (3) is true—but only *relative to a background framework partly constituted by the honor schema Igor has internalized*. By this I do not mean that Igor *believes* that (3) is true. I mean that it is *objectively* true as a claim about the “choice architecture” he confronts when deliberating about how to respond to the insult: given the choices made salient by the honor schema, issuing a challenge *is* the necessary means. Insofar as Igor internalized this schema, the possibility of elaborate practical jokes or stirring speeches about the evils of retribution are not live options, so issuing a challenge *really is* the necessary means to maintaining Igor’s honor. It is not logically, metaphysically, or nomologically necessary, of course. It is *practically* necessary in the sense defended by Williams: “what I recognize, when I conclude in deliberation that I cannot do a certain thing, is a certain incapacity of mine. I may be able to think of that course of action, but I cannot entertain it as a serious option.”²⁴ Thus we can construct Igor’s reason to issue a challenge in a quite straightforward way if we understand the scrutiny of our evaluative attitudes—our construction procedure—as constrained by the same choice-structuring assumptions as Igor’s own deliberations. This is the option I recommend.

The suggestion comes into better view with another example. Suppose Igor has internalized a queuing norm. One morning he finds himself wanting a cup

24 Williams, “Practical Necessity,” 128.

of tea on his way to an important appointment. He stops by a café and finds a line, one long enough that he will be late if he waits in it. Intuitively, the *pro tanto* reasons that Igor has arising from his valuing a cup of tea and getting to his appointment on time are in conflict, and so Igor needs to consider which value is more important to him. But how, exactly, does that conflict arise? The natural thought is that it is a nonnormative fact that Igor cannot both get tea and arrive at his appointment on time. But this is not strictly true. He *could* cut the line. Those queuing up are likely to be grumpy about it, but if he is insistent enough and offers elaborate justifications, he will probably prevail. So this is not enough to explain the alleged conflict. This appears to be another example of underdetermination.

One thought for overcoming it, akin to the option we considered above, would be to add that Igor values the coordination offered by the queuing norm or fears the reproach he would suffer for violating the norm. These attitudes would then give him reasons not to cut and so entail the conflict that we seem to find in his case. But it is not obvious that Igor *needs* to have these evaluative attitudes in order to face a conflict in this case. He could be skeptical about the value of the queuing norm. And he may be aloof enough that public reproach does not bother him. Even this version of Igor, it seems, faces a conflict between his competing reasons. If he has internalized the norm, then cutting is not an option, whether he likes the norm, is fearful of reproach, or not. He really *cannot* get tea and arrive on time. As in the previous example, our initial diagnosis of Igor's normative situation seems apt. Igor's evaluative attitudes really are in tension here, and this tension arises without mediation by auxiliary evaluative attitudes. But we can appreciate this tension only if we view Igor's practical point of view relative to certain background assumptions about the choices available to him—assumptions which, from Igor's point of view, are supplied by his internalization of the queuing norm.

In both of these cases I have suggested, first, that Street's constructivism cannot construct a reason that Igor seems, intuitively, to have and, second, that our intuitive way of understanding how Igor's reasons are constructed from his evaluative attitudes is correct, but only relative to a "background framework." These claims leave us with two questions. What are these "frameworks"? And what role, exactly, do they play in the construction of normative correctness?

I will not attempt a direct and comprehensive answer to these questions but instead approach them by way of analogy. The problem I have been saying Street's constructivism faces is similar to one confronted by philosophers of science. Scientific theories are underdetermined by the observations meant to support or refute them. For any set of observations, no matter how large, the class

of theories compatible with them—where we understand compatibility by the standards of logic alone—will be infinitely large. Only those theories that are logically inconsistent with one or more observation sentences will be ruled out, and that leaves an infinite remainder. This means that if we want to know which theory—or small set of theories—is supported by a set of observations, we are going to have to rely on something more than the meager constraints of logic.

Many philosophers of science have argued that, in practice, scientists overcome this underdetermination by relying on certain contingent but relatively well-entrenched background structures that shape how inquiry proceeds within particular research programs. These structures—Kuhn gave them their most famous name, “paradigms”—play a distinctive role in theory construction.²⁵ They define which sorts of problems are to be addressed in a given inquiry and sketch templates for solutions. They fix a representational scheme in which problems, observations, and explanations are to be couched. They precisify standards of theoretical coherence and articulate standard protocols for resolving violations of those standards, often by implicitly arranging theoretical goals into a hierarchy. They present experimental exemplars that guide empirical practice. They give operational glosses on vague theoretical values like simplicity and fecundity. This background work allows scientists to approach observations not as an undifferentiated heap that is compatible with an infinite number of theories, but as inputs into an articulate problem-solving apparatus.

If Kuhn’s descriptive claim about the role of paradigms is correct, then it seems inevitable that paradigms will have a coordinate effect on epistemic normativity—on what the scientist’s epistemic reasons are. For someone working within the paradigm of Skinnerian psychology, discovering some cognitive difference in a person is a reason to seek some corresponding conditioning mechanism in their environment and thus to design particular sorts of experiments crafted to discern such a mechanism. For someone working in a Newtonian paradigm, a certain style of solution to a mechanics problem will be epistemically appropriate—one assuming the laws of motion and done in the mathematicalized style of the *Principia*. For the phlogiston theorist, the cessation of combustion is a reason to think that the ambient air has been completely phlogisticated. This

25 Kuhn, *The Structure of Scientific Revolutions*. I use Kuhn as my example here because of his familiarity to most readers, but one can find devices that play similar roles in many places: the neo-Kantians’ (Hans Reichenbach, Michael Friedman) notion of the “relative” or “contingent” *a priori*, Imre Lakatos’s “research programs,” Michel Foucault’s “*savoir*,” Ian Hacking’s “historical” *a priori*, and Bas van Fraassen’s “stances.” There are important differences between these notions, but for my purposes they are not significant. See also Thompson, *Life and Action*, who emphasizes the role of life forms and practices in constituting normative standards, albeit by means rather different from those explored here.

connection between paradigm and epistemic normativity seems inescapable if we acknowledge that scientists' actual methodology—at least when it meets some minimal standard of sophistication—can affect what epistemic reasons they have. Without this concession, it would be hard to avoid the conclusion that the Skinnerian, the Newtonian, and the phlogiston theorist were not only mistaken in their picture of the world, but fundamentally irrational. (All this is to say nothing about the thorny question of whether the choice *between* different paradigms is subject to the standards of rationality, much less about the truth of the central claims of Skinnerian psychology or phlogiston theory.)

Of course, the *way* a paradigm conditions the scientist's epistemic reasons will be different from the way that her observations do. When determining what epistemic reasons arise in virtue of these observations, we do so by assuming features of the paradigm as fixed background conditions on the epistemic endeavor to which the reasons attach. That is, when working out what epistemic reasons a scientist has, we assume that she is solving a certain kind of problem, that certain features are desired in a solution, that certain sorts of experimental apparatuses and protocols are called for, that phenomena are to be represented in particular ways, that certain techniques are appropriate, and that a certain basic picture of the world obtains. In this way, the constraints a paradigm places on the determination of the scientist's epistemic reasons parallel those that it places on the same scientist's actual methods.

The notion of a "background framework" I have in mind is a rough practical analogue of these paradigms, and the constructivism I am proposing accords them analogous normative significance. Like Kuhnian paradigms, these background frameworks are not grounded in the demands of reason *as such*, and one can *in principle* be abandoned in favor of another. They are nonetheless relatively stable fixtures of practical reasoning that perform several crucial tasks that help us overcome the underdetermination of normative correctness by evaluative attitudes. They define a practical problem space by structuring an agent's choice architecture. They fix a scheme in which both values and practical problems are represented. They lay down standards of coherence for evaluative attitudes (ones stronger than those imposed by the constitutive standards of valuing). They specify protocols for resolving incoherence, often by implicitly arranging values into a hierarchy. They offer exemplars of practical problem-solving that guide agents' deliberations. They give operational glosses on vague values like courage and justice. The constructivism I am proposing as a remedy to the underdetermination of Street's is distinguished by the claim that background frameworks so understood also constrain the construction of normative correctness, and are, for that reason, normatively significant.

The crucial question here is: Constrained how? The idea must be to see the problem of what is normatively correct for an agent as structured in the same way as the problem the agent herself faces. For example, if a background framework structures an agent's choice architecture in a particular way, then, for the purposes of constructing what is normatively correct for her, we should take it as so structured. If it specifies particular *prima facie* evaluative tensions, then, for the purposes of normative construction, we should understand those values as standing in a *prima facie* tension. If it couches an agent's evaluative attitudes in particular concepts, our construction should do the same. If the framework privileges particular templates for solving practical problems, then our construction should privilege the same templates when deciding what reasons an agent has when facing such a problem.

I cannot offer a full delimiting of the range of constraints these frameworks may impose, but, as my examples suggest, I think these background frameworks very often are partly constituted by the further factors in the production of action that I identified in the previous section. And so those factors affect what is normatively correct for an agent by affecting their background structures of practical thought. For example, because Igor is embodied in a particular way, it is a feature of his background framework that the problem of seeing Manet's *Olympia* up close *just is* the problem of maneuvering his body into a particular region of space. As such, his wanting to see the painting entails, without the aid of auxiliary premises concerning his other evaluative attitudes and nonnormative facts, that he has a *pro tanto* reason to move his body in a suitable way.²⁶ By contrast, it is false that, for Igor, the problem of seeing *Olympia* up close *just is* the problem of flying to Paris and purchasing a ticket to the Musée d'Orsay—even if this is, in fact, the best way to see the painting—and so his wanting to see the painting does not entail that he has a reason to do these things, at least not without auxiliary premises. Similarly, because Igor has internalized an honor schema, an insult usually entails, often without any consideration of his other values, a reason to respond with a challenge, and because he has internalized a queuing norm, long lines and short time constitute a *prima facie* normative conflict. These are just the simplest and most direct ways that further factors can influence what is normatively correct for an agent, but if the claims of anthropologists like Clifford Geertz and Richard Shweder are correct, and there are

26 One way to put this is to say that the argument from “I want to see *Olympia*” to “I have a reason to occupy a particular region of space” is not an enthymeme. For a related argument, see Brandom, *Making It Explicit*.

profound socio-psychological differences between cultures, then we are likely to find equally profound normative differences.²⁷

Let me summarize the argument of this section. I began with a simple case that constructivists should acknowledge the normative significance of further factors in the production of action. I then noted two concerns with this idea: (i) the agent's "practical point of view" should only include those attitudes that reflect full-fledged normative judgment, and (ii) normative correctness is a matter of what is *really* entailed by these attitudes, not what an agent's parochial form of reasoning would *say* is entailed. These concerns naturally lead to a more austere constructivism, of which I took Street's view as an example. My reply, which has taken up most of this section, is that this spare constructivism faces a problem of normative underdetermination, one which can largely be attributed to its austerity. There are a few ways we can address this problem. One is to bite the bullet and concede the indeterminacy. The second is to be very liberal in attributing evaluative attitudes to agents. A third is to introduce a stronger conception of practical reason as such. None of these seem likely to succeed. The alternative I think is most plausible and most in the spirit of constructivism is to understand the construction of normative correctness from an agent's practical point of view as being guided not just by practical reason *as such*, but by a *background framework* that reflects contingent factors that color the agent's practical reasoning. Doing this gives those further factors an indirect but distinctive normative significance.

With this strategy we avoid concern (i) entirely, since we are not expanding our conception of the "practical point of view." As for (ii), I think we have good reason to reject it. It assumes that what is "really" entailed from a practical point of view is a matter of what is entailed by the universal, essential standards of practical reason as such (in Street's case, by the constitutive standards of valuing). But we should accept this assumption only if we think there is a single notion of practical entailment appropriate to all agents in all circumstances, and that this notion is up to the task of settling all the questions of normative correctness we expect to be settled. This is exactly what I think the underdetermination we have encountered gives us reason to doubt.

3. CONSTRUCTING MORAL REASONS

Constructivism is frequently criticized for failing to capture our considered judgments about normative correctness. The view produces "too few" reasons

27 I am thinking of Shweder's *Thinking through Cultures*, and Geertz's *The Interpretation of Cultures*.

because we can imagine cases in which we are inclined to say that an agent has a reason to do something even though she lacks the evaluative attitudes that would seem to be necessary for a construction of that reason from her practical point of view. She may, for example, be wholly and completely indifferent to the plight of a person on fire or to her own life, and thus have no reason, by the constructivist's lights, to douse the flames or take her medicine. And it produces "too many" reasons because agents might possess eccentric evaluative attitudes that would ground reasons that no one could really have: torture the innocent for amusement, count blades of grass, or be indifferent to her own pain so long as it is suffered on a Tuesday.

No single strategy is going to be adequate to this variety of objections. Kantian constructivists may have some success in answering them by appealing to the possibility that practical reason as such may ground moral and prudential reasons. Others will insist that constructivism is correct about these cases, and we will realize this if we better appreciate the psychology of the eccentrics involved.²⁸ There is a place for both strategies. What I want to suggest here is very limited: the constructivism I articulated in the previous section can contribute to this project of demonstrating the extensional adequacy of constructivism.

I focus on just one area: the ability of the constructivist to make good on our judgments about moral reasons. If an agent has the appropriate pro-moral evaluative attitudes—valuing the welfare of others, taking themselves to have a reason to respect rights, etc.—then we will be able to construct reasons for this agent to behave in morally appropriate ways. But some common opinions about moral reasons suggest that this is not enough. Many philosophers are attracted to moral rationalism, the view that everyone has moral reasons, whatever their evaluative attitudes, and that these reasons systematically override or exclude ones grounded in contingent interests. The Humean constructivist obviously cannot deliver on this. The Kantian constructivist can—sort of. If her arguments succeed, then everyone has an unconditional and overriding reason to follow whatever moral principle they have shown is legislated by practical reason itself, i.e., the Formula of Universal Law or Formula of Humanity. But the path from these principles to the reasons operative in particular cases is far from clear: Does my having an unconditional and overriding reason to act only on maxims I can will as universal laws entail that I have a similarly unconditional and overriding reason to leave a note when I dent your fender? Does my having an unconditional and overriding reason to treat humanity always as an end in itself and never as a mere means give me an overriding and unconditional reason to break one man's

28 E.g., Street, "In Defense of Future Tuesday Indifference," and Williams, "Internal Reasons and the Obscurity of Blame."

arm if it will save fifty others from death? The answer is unclear. This leaves the constructivist, even in the best-case scenario, with the problem of showing that those agents who lack the appropriate evaluative attitudes have moral reasons in the many cases in which the application of the Categorical Imperative is unclear (or, of course, arguing that they do not). Here is where I think the constructivism I just outlined can help the cause.

The key point will be about the normative effects of schemata. Some schemata will be connected to practices that we regard as purely conventional: queuing, fashion, and etiquette. But others will guide us in activities that seem morally significant: ones involving property, promises, others' bodies, or the categories of virtue. What is it that distinguishes moral schemata, customs, and social structures from ones of mere custom? A popular thought, which I adopt here, is that moral aspects of a social order are distinguished by being (constitutively) subject to certain forms of reflective criticism that mere customs are not. For example, Kurt Baier understands a "moral order" as

a social order which raises certain critical questions about its mores and which tends to modify them in light of the answers it gives to them. These questions therefore function as the society's own tests of soundness, that is, tests of the belief that certain directives contained in its mores (and possibly in those of other societies as well), and purporting to be moral directives and so to pass a certain appropriate test, really do pass it.²⁹

On this picture, there are two senses of morality: a particular social *practice* of policing and reforming other parts of the "social order," and the parts of the social order—rules, standards, norms—that purport to pass the distinctive tests imposed by this practice.

In an effective moral order, agents will be socialized so that they can successfully participate in that order, both by engaging in the critical activities of morality and by being guided by the results of that practice—by morality as their order finds it. And this means internalizing a schema that inclines them to behavior adhering to the prescriptions of that order's morality. It is easy enough to guess what this will entail. These agents will be inclined to represent situations with the import that morality prescribes: they will see an item as someone else's property, a speech act as a promise, a challenge as a test of fortitude. They will tend to interpret the actions of themselves and others according to the scheme imposed by morality: certain acts are thefts, promise breakings, acts of cowardice. They will have their choice architecture structured in a certain way: opportunities for theft may be systematically excluded from the range of options considered,

29 Baier, *The Rational and the Moral Order*, 212.

whatever action constitutes keeping a promise may be given default status, those actions that would be particularly vicious are silenced. And within the options that are considered, the social costs of contra-moral actions will be made particularly salient. This is the kind of psychological work that a schema must do in order to maintain agents' reproduction of the social order.

On the conception of constructivism sketched in the previous section, this *moralization* of an agent's background frameworks for practical thought will have consequences for what is normatively correct for them. What, exactly, these are obviously depends on the details of a given moral order and how it is internalized by the agents who enact it. That said, some general examples can be speculated about. An agent will not have reasons to perform actions that are excluded from consideration by her background framework. So if opportunities for petty theft are generally excluded from the choice architecture of agents socialized in a moral order, these agents have no reason to steal in these cases. On the other hand, if an option is bestowed with default status by an agent's background framework—e.g., because it would constitute keeping a promise—then this option will be *prima facie* choiceworthy for the agent. Other effects will be more indirect. The use of a particular scheme of representation that reflects the categories morality deems important—representing acts as theft, murder, promise breaking, etc.—and the association of these categories with particular evaluative and deontic vocabulary—“forbidden,” “permissible,” “generous,” etc.—will obviously not guarantee that an agent always has most reason to do what morality prescribes. But it will ensure that the demands of morality lie within the domain of practical thought for such agents. For these agents, adhering to the prescriptions of morality will always be an option to be taken or rejected. This will guarantee that morality is within the realm of things that agents *could* have reasons for. This may not sound like much, but it distinguishes morality from other normative systems, like dead religions or strange aesthetic practices, that lack a foothold in an agent's practical point of view. Finally, the increased salience of rewards and costs associated with pro- and contra-moral behavior will tend, in the long run, to foster evaluative attitudes that will ground reasons of moral compliance.

My claim is that morality will be normative in these and allied senses for agents who have internalized the schemata distinctive of social orders that are also moral in Baier's sense. What does this get us? Certainly not the strong version of moral rationalism I described a moment ago. Even for agents who have internalized a moral order, it will not guarantee that they have reasons to do what that morality advises (much less *decisive* reason). Instead, I have suggested more attenuated forms of normative “significance” that under some condi-

tions ground reasons for action. And even these limited effects will be fragile, since schemata are not deliberative straitjackets. An agent's internalization of a particular schema may silence the option of stealing my tea or wallet or make promise-keeping a deliberative default, but the right circumstances and a strong enough desire for tea or cash can cancel these effects. Indeed, professional burglars, faithless schemers, and full-fledged morality critics can condition themselves to neutralize the pro-moral effects of the common schema and in doing so remove these elements from their background framework of practical reasoning. And this is to say nothing of those who leave a given moral order altogether. The proposal falls short of full-blooded moral rationalism in another respect as well. The conception of morality that has normative significance for agents on this picture is tied to a particular social order. This is not to say that it is arbitrary, since it must pass the tests of validity that are constitutive of the difference between morality and mere custom. (Something that Baier discusses at length.³⁰) But there is room for significant variation from one community to another, so we are not going to find reasons for any objective conception of morality.

Even with these concessions, I think the possibility outlined here has the potential to improve the fortunes of constructivists in trying to produce a satisfactory account of the moral reasons that agents have. First, by locating some of the normative significance of morality in agents' background framework of practical thought, we are able to explain why, in certain instances, the normative force of morality is unconditioned by agents' particular evaluative attitudes. Concomitantly, we can explain how morality can, sometimes, exclude or override potential reasons for contra-moral action, e.g., by excluding those options from the agent's choice architecture. In this respect it is an improvement on what the Humean constructivist can do. Second, by showing how the conceptual categories of a particular moral order gain entry into agents' practical thought, we are able to explain why agents have the reasons we intuitively judge them to have in concrete, contextualized cases—why they have no reason to commit some petty theft—and not just a reason to follow a highly abstract formulation of the moral law. In this respect, it is a useful supplement to extant versions of Kantian constructivism.

If constructivists are going to offer an account of normativity that comports with our considered judgments, they will have to fight on many fronts. They will have to show how more can be grounded in our evaluative attitudes than we might have thought. They may very well have to show that the use of practical reason as such commits us to a general moral principle. They will have to show that some of our initial judgments about reasons are unfounded. What I

30 Baier, *The Rational and the Moral Order*, 197–202.

am claiming here is not that the liberal conception of constructivism sketched in the previous section is a panacea for these problems, but that it can play an important role in allowing constructivists to capture more of the nuance of the normative world.

4. PRACTICAL REASON NOT AS SUCH

In closing, I want to briefly suggest that my argument here can be understood as a case for there being such a thing as practical reason *not as such*—for the legitimacy of talking about the ways in which practical reason, and not just its objects, can be shaped by the contingent features of an agent's situation and constitution.

There are two issues at stake in this suggestion. One is descriptive: Does positing such a thing as practical reason not as such help us explain human behavior? My argument in the first section is devoted to showing that it does. I argued that further factors shape an agent's practical outlook and structure her deliberations in ways that go well beyond our paradigms of valuing. In particular, they condition an agent's capacity for practical reasoning without being features of practical reasoning as such in a fashion analogous to the role of individual grammars in linguistic reasoning.

The second is normative: Does this faculty help us explain what reasons an agent has, what is good for her, and so on? In the second section I argued that, insofar as we have roughly constructivist metaethical scruples, it is appropriate to recognize these factors as making a distinctive contribution to normative correctness. I characterized this contribution as one of background conditions on the construction procedure, but we could just as well understand the construction of normative correctness as undertaken by a single capacity of practical reason that encompasses both this "background framework" and practical reasoning "as such." We could, that is, say that differences among agents in further factors actually carve out different *styles* or *forms* of practical reason—and different standards of correctness—just as the deep differences in scientific method reflect, in Ian Hacking's terms, different "styles" of theoretical reasoning.³¹

Of course, to make good on this project we would need much more precise criteria of individuation than I have provided here. Nonetheless, I think these results give us some initial reason to countenance talk of practical reason *not as such*, both as part of the project of explaining action and understanding the grounds of normative correctness. Indeed, unless we are keen on the question of practical reasoning's *essential* commitments, the division between practical

31 Hacking, "Language, Truth, and Reason."

reason “as such” and other features of agents’ practical thinking can look highly artificial.

Humean and Kantian constructivists *are* interested in this question, since their principal disagreement is over whether there are any substantive judgments that are normatively correct for all possible rational creatures. To them, the constructivism I have defended here might look like no more than a modest expansion of the Humean program, since it has no pretensions to establishing the universality of substantive normative judgments. But this is not the only important question we can ask about the etiology of normative facts. It matters whether the reasons we have to adhere to some set of norms—be they a particular conception of morality, a religious creed, or a racist ideology—are rooted in widespread but nevertheless personal evaluative attitudes, or in contingent features of practical thought traceable to agents’ socialization or embodiment. This in turn makes a difference to anyone taking up a critical perspective on these norms. Suppose a person comes to question the morality, religion, or racism she hitherto thought was normatively appropriate for her. How should she respond? If the normative correctness reflects her evaluative attitudes and no more, then she should reevaluate (which may sometimes require substantial effort). But if they are grounded in her internalization of a social schema, then a more thoroughgoing transformation is necessary. She must retrain herself in the use of practical reason to extirpate the suspect influences, and, in some cases, try to overturn whatever social structures introduce and sustain those influences.³²

The general point is that it is often not enough to understand *that* a certain judgment is normatively correct for us, nor that it is *contingently* or *necessarily* so. We must also understand *why* it is correct, *which* aspects of our constitution and situation contribute to that correctness, and *how* they so contribute. Appreciating the normative significance of those features that constitute practical reason *not as such* is a first step in that direction.³³

Dartmouth College
kenneth.e.walden@dartmouth.edu

32 On this distinction in the case of anti-racism in particular, see Shelby, “Racism, Moralism, and Social Criticism,” and Haslanger, “Racism, Ideology, and Social Movements.”

33 I am grateful to many people for comments and helpful discussion. The following list is just some of these people; I apologize for the inevitable omissions: Richard Holton, Julia Markovits, Alice Phillips Walden, David Plunkett, Paulina Sliwa, Sharon Street, an audience at the Madison Metaethics Workshop, and two anonymous referees.

REFERENCES

- Baier, Kurt. *The Rational and the Moral Order*. Chicago: Open Court, 1995.
- Bicchieri, Cristina. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press, 2006.
- Bicchieri, Cristina, and Yoshitaki Fukui. "The Great Illusion: Ignorance, Informational Cascades and the Persistence of Unpopular Norms." *Business Ethics Quarterly* 9, no. 1 (January 1999): 127–55.
- Bourdieu, Pierre. *The Logic of Practice*. Stanford: Stanford University Press, 1992.
- Brandom, Robert. *Making It Explicit*. Cambridge, MA: Harvard University Press, 1994.
- Frankfurt, Harry. "Autonomy, Necessity, and Love." In *Necessity, Volition, and Love*, 129–41. New York: Cambridge University Press, 1998.
- Friedrich, Paul. "Sanity and the Myth of Honor: The Problem of Achilles." *Ethos* 5, no. 3 (Fall 1973) 281–305.
- Geertz, Clifford. *The Interpretation of Cultures*. New York: Basic Books, 1977.
- Hacking, Ian. "Language, Truth, and Reason." In *Rationality and Relativism*, edited by Martin Hollis and Steven Lukes, 48–66. Cambridge, MA: MIT Press, 1982.
- Haslanger, Sally. "Racism, Ideology, and Social Movements." *Res Philosophica* 94, no. 1 (January 2017): 1–22.
- . "What Is a (Social) Structural Explanation?" *Philosophical Studies* 173, no. 1 (January 2016): 113–30.
- Hlobil, Ulf. "Social Norms and Unthinkable Options." *Synthese* 193, no. 8 (August 2016): 2519–37.
- Korsgaard, Christine. *Sources of Normativity*. New York: Cambridge University Press, 1996.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
- Lenman, James. "Humean Constructivism in Moral Theory." In *Oxford Studies in Metaethics*, vol. 5, edited by Russ Shafer-Landau, 175–94. New York: Oxford University Press, 2010.
- MacIntyre, Alasdair. *Is Patriotism a Virtue?* Lindley Lectures, vol. 23. Lawrence, KS: University of Kansas, 1984.
- Markovits, Julia. *Moral Reason*. New York: Oxford University Press, 2014.
- Pettit, Philip, and Michael Smith. "Backgrounding Desire." In *Mind, Morality, and Explanation: Selected Collaborations*, edited by Frank Jackson, Philip Pettit, and Michael Smith, 269–94. New York: Oxford University Press, 2004.

- Shelby, Tommie. "Racism, Moralism, and Social Criticism." *DuBois Review* 11, no. 1 (Spring 2014): 57–74.
- Shweder, Richard. *Thinking through Cultures*. Cambridge, MA: Harvard University Press, 1991.
- Street, Sharon. "Constructivism about Reasons." In *Oxford Studies in Metaethics*, vol. 3, edited by Russ Shafer-Landau, 207–46. New York: Oxford University Press, 2008.
- . "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues* 19, no. 1 (October 2009): 273–98.
- . "What Is Constructivism in Ethics and Metaethics?" *Philosophy Compass* 5, no. 5 (May 2010): 363–84.
- Taylor, Charles. "Lichtung or Lebensform: Parallels between Heidegger and Wittgenstein." In *Philosophical Arguments*, 61–78. Cambridge, MA: Harvard University Press, 1995.
- . "The Validity of Transcendental Arguments." In *Philosophical Arguments*, 20–33. Cambridge, MA: Harvard University Press, 1995.
- Thompson, Michael. *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge, MA: Harvard University Press, 2008.
- Williams, Bernard. "Internal and External Reasons." In *Moral Luck*, 101–13. New York: Cambridge University Press, 1981.
- . "Internal Reasons and the Obscurity of Blame." In *Making Sense of Humanity and Other Philosophical Papers*, 35–45. New York: Cambridge University Press, 1995.
- . "Practical Necessity." In *Moral Luck*, 124–31. New York: Cambridge University Press, 1981.

IN OR OUT?
ON BENEVOLENT ABSOLUTISMS
IN *THE LAW OF PEOPLES*

Robert Huseby

BENEVOLENT ABSOLUTISMS occupy a rather unclear position in Rawls's *The Law of Peoples*. On the one hand, these states, because they are not well ordered, do not belong to the Society of Peoples, which consists of those societies that are both law abiding and well ordered. On the other hand, unlike other societies that are either not well ordered or not law abiding (or both), benevolent absolutisms are not to be assisted or sanctioned into becoming well ordered. Given Rawls's aim of expanding the Society of Peoples for the benefit of lasting peace and stability, this situation seems wanting. In light of this, I argue that *The Law of Peoples* should be altered in order to clarify the theoretical status of benevolent absolutisms, and I discuss alternative ways of doing so. First, I consider including these states into the Society of Peoples. This solution is problematic in part because it would implausibly strain the notion of liberal tolerance. Second, I consider merging the two criteria for membership in the Society of Peoples—well-orderedness and adherence to the Law of Peoples—by making the latter a part of the former. As it turns out, this does not solve many problems, and I therefore suggest the further move of including a crucial aspect of the well-orderedness criterion into the very *conception of human rights* contained in the Law of Peoples. This, I argue, does clarify the position of benevolent absolutisms. These states no longer meet one of two criteria for inclusion into the Society of Peoples. They now fail the only criterion there is. Further, making this aspect of well-orderedness a part of the conception of human rights opens the possibility of subjecting benevolent absolutisms to sanctions. The reason is that, on Rawls's view, respecting human rights excludes the imposition of justified sanctions.

1. INTRODUCTION

A central theme in the later writings of John Rawls is how individuals and

groups with incompatible worldviews can live together in peace and stability. The question is crucial in both the domestic and international realms. In *Political Liberalism*, Rawls presents a notion of legitimacy that aims to facilitate stable and peaceful cooperation between adherents to a range of different and incompatible comprehensive doctrines. He claims that, in liberal constitutional democracies marked by reasonable pluralism, it would be unreasonable for citizens to insist that the basic structure of society should be organized in light of their own particular doctrine. Rather, all citizens should accept some ground rules for the organization of the political sphere that all reasonable persons, regardless of their comprehensive doctrines, can endorse.¹

In *The Law of Peoples*, Rawls extends some of his core ideas on reasonably just constitutional democracies to the international sphere.² More precisely, he formulates the principles that ought to guide liberal peoples' foreign policy.³ Just as members of comprehensive doctrines within a domestic society should accept political liberalism, so should peoples accept an ideal international law—the Law of Peoples—that is to regulate international cooperation and interaction. This is a contested claim from a liberal point of view, since adherence to the Law of Peoples does not require internally liberal institutions.

For theoretical purposes, Rawls proposes a (non-exhaustive) ideal-typical categorization of societies. The first type consists of reasonable liberal peoples, which are stable democracies organized in light of a liberal political conception of justice, of the kind outlined in *Political Liberalism*. The second type is decent peoples. These could take various forms, but Rawls mainly discusses a type that is hierarchically organized according to a religious doctrine. The third type is so-called burdened societies, “whose historical, social, and economic circumstances make their achieving a well-ordered regime, whether liberal or decent, difficult if not impossible.”⁴ These societies are unable, rather than unwilling, to accept the demands of the Law of Peoples. The fourth is outlaw states. These are aggressive, and they do not honor either human rights or the Law of Peoples more generally.⁵ Finally, there are benevolent absolutisms. These societies

1 Rawls, *Political Liberalism*, 12.

2 Rawls, *The Law of Peoples*, 9.

3 Rawls uses the term “peoples” as opposed to “states” for societies that are law abiding and well ordered.

4 Rawls, *The Law of Peoples*, 5.

5 Rawls, *The Law of Peoples*, 90.

respect human rights, but deny their members any meaningful role in political decision-making.⁶ They are therefore not well ordered.⁷

Only reasonable liberal peoples and decent hierarchical peoples fulfill the criteria for admission into what Rawls refers to as the Society of Peoples. First, these peoples are law abiding in that they accept the principles of the Law of Peoples, and second, they are *well ordered*. It appears that both criteria are necessary, and jointly sufficient, for inclusion into the Society of Peoples.⁸ The three latter types of states—burdened societies, outlaw states, and benevolent absolutisms—are treated in the nonideal part of the theory, covering cases of non-compliance and unfavorable conditions. It is noteworthy, however, that while outlaw states and burdened societies are discussed in detail, very little is said about benevolent absolutisms.

In this paper, I will be primarily concerned with these benevolent absolutisms and their apparently uneasy status within Rawls's theory. On the one hand, as noted, they fall outside the Society of Peoples; on the other hand, liberal and decent hierarchical societies are neither to assist nor sanction them because they are peaceful and respect human rights. Given the aims of the theory, which is to expand a peaceful Society of Peoples, this is unfortunate. I argue that *The Law of Peoples* should be altered in order to clarify the status of these states, and discuss different ways of doing so.

While *The Law of Peoples* has been met with substantial criticism (and defense), there is a need to see how far his theory can be amended and adjusted in light of the criticism before a final assessment is made. The present paper is a contribution to that effort. If I am right that benevolent absolutisms occupy an awkward position in the theory, one pertinent question to ask is whether the theory can be slightly amended so as to remove this awkwardness. This is important also from a practical point of view, since we need a coherent and principled view on how liberal peoples should respond to various kinds of non-liberal states. *The Law of Peoples* offers such guidance when it comes to decent hierarchical societies, burdened societies, and outlaw states, but more needs to be said about benevolent absolutisms, or so I argue.⁹

6 Rawls, *The Law of Peoples*, 4.

7 Rawls, *The Law of Peoples*, 4, 63, 92.

8 Rawls sometimes seems to suggest that respecting the Law of Peoples is sufficient. See for instance Rawls, *The Law of Peoples*, 3. It is clear from his remarks on benevolent absolutisms, however, that well-orderedness is required as well.

9 For related contributions, see Maffettone, *The Coherence and Defensibility of Rawls's Law of Peoples*, "Benevolent Absolutisms, Incentives, and Rawls's *The Law of Peoples*," and "Should We Tolerate Benevolent Absolutisms?"

On a methodological note, I should emphasize that my goal is not to argue in favor of an ideal system of international law as such. Rather, I start out from an apparent incongruence in the Rawlsian way of dealing with benevolent absolutisms, and propose a strategy for dealing with this from a perspective internal to that theory.¹⁰

2. THE LAW OF PEOPLES, THE SOCIETY OF PEOPLES, AND WELL-ORDEREDNESS

In Rawls's contractual theory, the Law of Peoples is first chosen by representatives of liberal peoples in a hypothetical contract situation under a (suitably tailored) veil of ignorance.¹¹ The result is, unsurprisingly, broadly liberal, because liberal peoples choose the law that they ideally think should guide international interaction. However, this ideal is informed by a—somewhat controversial—notion of tolerance. Due in part to this idea of tolerance, the Law of Peoples will be acceptable to some non-liberal peoples as well. Specifically, Rawls argues that representatives of decent hierarchical peoples will adopt the *same* law in their own (subsequent) hypothetical contract.¹² The Law of Peoples consists of eight principles:

1. Peoples are free and independent, and their freedom and independence are to be respected by other peoples.
2. Peoples are to observe treaties and undertakings.
3. Peoples are equal and are parties to the agreements that bind them.
4. Peoples are to observe a duty of non-intervention.
5. Peoples have the right of self-defense but no right to instigate war for reasons other than self-defense.
6. Peoples are to honor human rights.
7. Peoples are to observe certain specified restrictions in the conduct of war.
8. Peoples have a duty to assist other peoples living under unfavorable conditions that prevent their having a just or decent political and social regime.¹³

10 For more critical approaches, see the references in note 41.

11 The parties do not know their country's size, population, strength, possession of natural resources, or level of economic development. On the other hand, "they do know that reasonably favorable conditions obtain that make constitutional democracy possible—since they represent liberal societies" (Rawls, *The Law of Peoples*, 32–33). Liberal peoples have already chosen their *domestic* principles of justice in domestic original positions (Rawls, *A Theory of Justice*).

12 Rawls, *The Law of Peoples*, 64.

13 Rawls, *The Law of Peoples*, 37.

The Society of Peoples is Rawls's term for an idealized, peaceful, and stable association of peoples that are all internally well ordered, and share a desire to respect and uphold the Law of Peoples. The members of the Society of Peoples have a general aim of making other states respect the Law of Peoples as well. Burdened societies are to be *assisted* (because they are *unable* to become well ordered on their own), while outlaw states may be *sanctioned* (because they are *unwilling* to become law abiding) by way of diplomatic, economic, or military means.

The overarching goal is to secure lasting peace and stability among peoples. Rawls states that it is a "basic characteristic of well-ordered peoples that they wish to live in a world in which all peoples accept and follow (the ideal of the) Law of Peoples."¹⁴ He also writes that "it is characteristic of liberal and decent peoples that they seek to live in a world in which all peoples have a *well-ordered regime*."¹⁵ The natural way of achieving this goal is to work toward bringing more and more states into the Society of Peoples. Extending the Society of Peoples is likely to benefit all societies, since well-ordered and law-abiding societies are (by definition) not aggressive. Any extension will more particularly benefit those peoples that become members of the Society of Peoples, because membership presupposes just or decent, that is, well-ordered, domestic institutions.¹⁶

Decent hierarchical peoples are well ordered partly in virtue of allowing their members a *meaningful role in making political decisions*. Allowing members such a role is of fundamental importance for Rawls. This is a feature that benevolent absolutisms lack, and this is the *main reason* why they cannot be considered well ordered or candidates for membership in the Society of Peoples. Decent hierarchical peoples are defined by two criteria. These are of interest here as they concern the idea of well-orderedness.

1. First, the society does not have aggressive aims, and it recognizes that it must gain its legitimate ends through ... ways of peace.
 - 2a. A decent hierarchical people's system of law ... secures ... human rights.
 - 2b. A decent people's system of law must ... impose *bona fide* moral duties and obligations (distinct from human rights) on all persons within the people's territory.
 - 2c. There must be a sincere and not unreasonable belief on the part of judges and other officials ... that the law is indeed guided by a common good idea of justice.¹⁷

14 Rawls, *The Law of Peoples*, 89.

15 Rawls, *The Law of Peoples*, 113, emphasis added.

16 For a discussion, see Huseby, "John Rawls and Climate Justice."

17 Rawls, *The Law of Peoples*, 64–67.

In order to meet criteria 2b and 2c, Rawls argues that decent hierarchical societies must have some form of “decent consultation hierarchy.”¹⁸

In political decisions a decent consultation hierarchy allows an opportunity for different voices to be heard. . . . Persons as members of associations, corporations, and estates have the right to at some point . . . (often at the stage of selecting a group’s representatives) to express political dissent, and the government has an obligation to take a group’s dissent seriously and to give a conscientious reply.¹⁹

As noted, Rawls holds that benevolent absolutisms, unlike decent hierarchical societies, do not give their members a meaningful role in political decision-making, and are *therefore* not well ordered.²⁰ In line with this, I take it that benevolent absolutisms could become well ordered by securing for their members such meaningful political participation. This would, as Rawls suggests, secure that the law imposes *bona fide* moral duties and obligations on all participants, and that judges and officials are sincere in their belief that the law is guided by a common good conception of justice.²¹ In other words, if a benevolent absolutism instituted a decent consultation hierarchy, it would both secure meaningful political participation for its members *and* go from non-well ordered to well ordered. The reason is that doing so would make sure it fulfills 2c and 2b.

18 Rawls, *The Law of Peoples*, 71.

19 Rawls, *The Law of Peoples*, 72.

20 Rawls, *The Law of Peoples*, 4, 63, 92, emphasis added. See Maffettone, *The Coherence and Defensibility of Rawls’s Law of Peoples*, 24–56, for discussion of why political participation is so crucial to Rawls’s understanding of well-orderedness. My understanding of the importance of political participation in relation to benevolent absolutisms seems to be more in line with Neufeld, “Liberal Foreign Policy and the Ideal of Fair Social Cooperation,” 296.

21 Rawls, *The Law of Peoples*, 92. Rawls draws in his discussion of these issues on Philip Soper’s theory of law (Soper, *A Theory of Law*). Arguably, however, there could be *some bona fide* moral duties imposed by law in benevolent absolutisms, even if we assume that Rawls’s Soper-inspired theory of political obligation is overall correct. As pointed out by Estlund, legitimate authority, and subsequent obligation to obey, can arise in different ways, from different sources (*Democratic Authority*). Thus, it is possible that the law, even in benevolent absolutisms, can impose *bona fide* moral duties on citizens for reasons such as (a) the law actually protects valuable moral rights (apart from human rights, which clearly also impose moral duties on citizens), and (b) law-abidingness, to some extent, is necessary to preserve societal stability over time. There could be further reasons as well. However, even if the law in benevolent absolutisms *can* impose *bona fide* moral duties (*contra* Rawls), it can still be true that the system cannot impose such duties with the generality and robustness necessary for achieving decency or meaningful political participation. Generally, I suppose systems of law can fall short of decency in many different ways and to many different degrees. I am very grateful to an anonymous reviewer for helpful comments on this issue.

Benevolent absolutisms already meet 1 and 2a, since they are not aggressive and they respect human rights.²²

3. BENEVOLENT ABSOLUTISMS

Quite a lot has been written about *The Law of Peoples*. Less, however, has been said specifically about the “curious case of benevolent absolutisms.”²³ These societies occupy an uneasy position in the theory. Unlike burdened societies, they are not unable to become well ordered and members of the Society of Peoples, and there is thus no need to assist them. And unlike outlaw states, they are not to be subject to sanctions, since they, like decent societies, are nonaggressive and respect human rights.²⁴ Nonetheless, these societies must be unwilling on some level (since they are not unable), to become well ordered. In any case, because they are not well ordered, they are not eligible as members of the Society of Peoples, and the international community, moreover, has no means with which to put pressure on them.

3.1. *Can Benevolent Absolutisms Respect the Law of Peoples?*

The role of benevolent absolutisms is somewhat peculiar, since there is nothing to say that these societies could not accept the Law of Peoples in its entirety.²⁵ They respect and honor human rights, which is the only principle of the Law of Peoples that explicitly addresses the internal organization of participating states.²⁶ Further, there is no reason to think that benevolent absolutisms would be opposed to the principles concerning self-determination and noninterven-

22 Well-orderedness is given a stricter interpretation in Rawls’s *domestic* theory of justice. Liberal constitutional democracies are well ordered when “everyone accepts, and knows that everyone else accepts, the very same political conception of justice” (Rawls, *Justice as Fairness*, 8). Further, the basic structure of such societies is believed by most citizens to meet the standards of justice they affirm (*Justice as Fairness*, 8). Lastly, “citizens have a normally effective sense of justice . . . that enables them to understand and apply . . . the principles of justice” (*Justice as Fairness*, 8). Decent hierarchical societies are not well ordered in light of these criteria.

23 This phrase is Maffettone’s, who is one of the few to offer substantive treatments of benevolent absolutisms. See Maffettone, *The Coherence and Defensibility of Rawls’s Law of Peoples*, 142–53, “Benevolent Absolutisms, Incentives and Rawls’s *The Law of Peoples*,” and “Should We Tolerate Benevolent Absolutisms?” See also Neufeld, “Liberal Foreign Policy and the Ideal of Fair Social Cooperation”; Reidy, “Human Rights and Liberal Toleration”; and Riker, “The Democratic Peace Is Not Democratic.”

24 Rawls, *The Law of Peoples*, 80, 92.

25 Maffettone, *The Coherence and Defensibility of Rawls’s Law of Peoples*, 143.

26 Rawls, *The Law of Peoples*, 92.

tion. It is also likely that such states would be sympathetic to the principles pertaining to treaties, international cooperation, and the conduct of war. Thus, it appears that benevolent absolutisms most likely accept the Law of Peoples.

One might ask whether benevolent absolutisms can *really* accept and honor the Law of Peoples, however, since these states are not peoples at all, precisely because members are denied a meaningful political role. As such, these states do not constitute corporate agents.²⁷ However, these states, by their *leaders*, can, in one sense at least, accept and honor the Law of Peoples. They can, for instance, abide by its statutes and forgo any violations of them. Suppose there are two states, one well ordered and one not. Even though only the well-ordered state properly acts on behalf of its people, both states, however constituted, act meaningfully in the world. Both states can intervene or not into other states, both can sign treaties, and both can uphold laws internally. Thus, even though the Law of Peoples is intended to regulate the interactions between one type of international actor—peoples—it seems perfectly possible for another kind of actor, benevolent absolutisms, to honor, comply with, and accept this law. Thus, whether or not benevolent absolutisms accept the Law of Peoples does not hinge on whether they constitute a suitable corporate agent, but on whether the regime commits to *actually* honoring and accepting it.

This, moreover, seems compatible with Rawls's view. Even though human rights do not constitute the whole of the Law of Peoples, they do constitute a part of it, and if benevolent absolutisms can "respect" human rights, they can respect (and honor) the other principles as well. According to Rawls, "while a benevolent absolutism does respect and honor human rights, it is not a well-ordered society, since it does not give its members a meaningful role in making political decisions."²⁸ This quote strongly suggests that benevolent absolutisms can indeed accept, in a relevant sense, international law. Further, if benevolent absolutisms were unable, due to their lack of well-orderedness, to respect international law, they would also, for that *same* reason, fail to respect human rights, which would make them vulnerable to justified sanctions.²⁹ Notice also that when Rawls discusses outlaw states, he says that they are *unwilling* to abide by the Law of Peoples, and not that it is in principle impossible for them to do so because they are not well ordered and do not constitute an appropriate kind of corporate agent.³⁰

27 Consequently, Pettit argues that societies that are not well ordered, have "no standing under the law of peoples" ("Rawls's Peoples," 43).

28 Rawls, *The Law of Peoples*, 92.

29 I say more about the connection between human rights and sanctions below.

30 Rawls, *The Law of Peoples*, 90. Burdened societies, on the other hand, are unable to comply

This interpretation is further strengthened by the fact that Rawls holds that (in addition to well-ordered societies), “any society that follows and honors a reasonably just Law of Peoples” has a right to self-defense, *and* that benevolent absolutisms, specifically, have a right to self-defense.³¹ This does not strictly imply that benevolent absolutisms honor the Law of Peoples, but, in my view, it strongly indicates as much, since the quote does imply that even non-well-ordered peoples can follow and honor the Law of Peoples, and benevolent absolutisms are seemingly as close to well ordered as can be. There is a distinction, then, between a leadership properly representing its people and a leadership honoring international law. Thus, benevolent absolutisms can honor the Law of Peoples even if they do not constitute a people.

One might perhaps question whether these states would accept the last principle, concerning the duty of assistance. Not because there is any particular reason to suppose that benevolent absolutisms would be averse to assistance among societies, but because the goal of the assistance is to enable burdened societies to become well ordered. However, the principle as it is stated seems compatible with different motivations, and it is clearly conceivable that benevolent absolutisms could accept even this principle, though perhaps with a different motivation than hierarchical and liberal societies.

3.2. *The Theoretical Purpose of Benevolent Absolutisms*

Rawls’s typology of states is not perfectly systematic, and one might wonder what theoretical purpose benevolent absolutisms serve. Why, for instance, are they defined as nonaggressive? This is not entirely clear, but it seems that if they had been defined as aggressive they would have been subject to sanctions, though not for the same reason as outlaw states. There is room in Rawls typology (which as noted is non-exhaustive) for two kinds of semi-benevolent (or semi-outlaw) absolutisms. The first would be like outlaw states in that they would be externally aggressive, but like benevolent absolutisms in that they would respect human rights internally. The other would be like outlaw states in that they would violate human rights internally, but like benevolent absolutisms in that they would be externally nonaggressive. Neither kind would, then, be well ordered nor give their citizens a meaningful political role. Both kinds of society would potentially be subject to sanctions, either because of their external

with the Law of Peoples, but this is due to social, historical, and economic factors, not a (possible) lack of political participation (Rawls, *The Law of Peoples*, 4).

31 Rawls, *The Law of Peoples*, 91, emphasis added, and 92.

aggression or because of their human rights violations. Outlaw states are potentially subject to sanctions on both counts.³²

Now, the reason why Rawls assumes that benevolent absolutisms are non-aggressive, presumably, is to carve out a conceptual space between decent societies and outlaw states. Outlaw states are primarily defined in virtue of violating human rights internally and being aggressive externally. Benevolent absolutisms lack both these features, but are not well ordered, because they do not give their citizens a meaningful role in political decision-making. And such a meaningful role is one of the most important features in light of which decent states are decent.

A similar contrast can be made between benevolent absolutisms and burdened societies. Burdened societies are not well ordered, and because they lack the capacity to be so, they are consequently not included in the Society of Peoples, and there is a duty to assist them. Benevolent absolutisms presumably have the capacity to become well ordered, since there is no duty to assist them.

Benevolent absolutisms, then, is a useful category because they fill out (some of) the conceptual space between the group of states that are part of the Society of Peoples and the group of states that are subject to sanctions or targets of a duty of assistance. In addition, benevolent absolutisms arguably have empirical counterparts (see below). One could, of course, imagine states that give their citizens a meaningful role in political decision-making but fail to be well ordered for other reasons. However, as indicated below, Rawls seems to employ a less demanding conception of well-orderedness in *The Law of Peoples* than in *Justice as Fairness*. Hence, benevolent absolutisms fit into the space between decent societies, outlaw states, and burdened societies in a theoretically suitable and illustrative way (though I will, as noted, question whether there should be such a space).

This theoretical position also helps explain why benevolent absolutisms have the right to military self-defense. According to Rawls, all societies that honor ideal international law have the right to defend themselves in the face of external aggression.³³ It makes sense that benevolent absolutisms have the right to self-defense despite not being well ordered. As noted, they are, on Rawls's view, immune to sanctions, in virtue of being peaceful and respecting human rights.³⁴ If they had not been immune to sanctions, decent and liberal states might have

32 Rawls claims that respecting human rights immunizes a state against sanctions (*The Law of Peoples*, 80). It is clear, however, that external aggression can trigger sanctions too, whether or not such aggression typically violates the human rights of the victims of the aggression.

33 Rawls, *The Law of Peoples*, 91.

34 Rawls, *The Law of Peoples*, 80.

had grounds for sanctions and, in extreme cases, even military intervention. If so, it would have made sense to claim that benevolent absolutisms should not have a right to self-defense. Otherwise there could be situations in which a benevolent absolutism could have a right to self-defense against a justified intervention. As it is, the *only* kinds of states that pose a military threat to benevolent absolutisms are outlaw states. It would be highly implausible to hold that benevolent absolutisms do not have the right to defend themselves against armed aggression from outlaw states.

3.3. *Do Benevolent Absolutisms Only Have Theoretical Interest?*

One might also ask whether benevolent absolutisms are of more than theoretical interest. I think the answer is yes. First, some historical societies can with some justification be labeled benevolent absolutisms (or benevolent despotisms or enlightened absolutisms).³⁵ Further, many current states share, or have shared, important characteristics with benevolent absolutisms.³⁶ These examples are only suggestive, for two main reasons. First, because they capture only parts of the features of benevolent absolutism and, second, because benevolent absolutisms are ideal types and we cannot necessarily expect to find many clear-cut examples in the real world. Nevertheless, the examples do suggest, quite clearly, that benevolent absolutisms are of more than theoretical interest. It is important, therefore, that liberal (and decent) societies have a principled basis on which to interact with such states. Further, regardless of how many states that today fit the description, there is no telling what the future might hold. Some democracies might regress and some outlaw states (or burdened societies) might progress in such a way that they for some period of time become peaceful and human rights respecting but without providing meaningful political participation for their inhabitants.

35 So-called enlightened despotism is associated with seventeenth-century monarchs in Prussia, Russia, and Austria.

36 These include: Swaziland (1976, 1978, 1980–82, 1986, 1988, 1992), Saudi Arabia (1976, 1978), Kuwait (1976–79, 1988), Qatar (1978–81, 1987–90, 1992–99, 2000–06), Oman (1976–77, 1988, 1990), Bhutan, (1982–86, 1988–89, 1996, 1998, 2000–03), Argentina (1972), Panama (1985–86), Liberia (1974), Thailand (1973), Honduras (1973–79), and Fiji (2000). The first six of these countries are the result of combining a high degree of power concentration (Polity = -10; Marshall, Gurr, and Jaggers, *Polity IV Project*) with rare instances of human rights violations (Political Terror Scale = 1; Wood and Gibney, “The Political Terror Scale (PTS)”). The next six are the result of combining the near absence of any political rights (PR = 7) with a moderate level of civil liberties (CL = 3) in the Freedom House index (Freedom House, *Freedom in the World 2016*). I am indebted to Håvard Strand for very generous help with this issue. Note also that Maffettone suggests Brunei as a possible example of a benevolent absolutism (Maffettone, *The Coherence and Defensibility of Rawls’s Law of Peoples*).

Further, even if benevolent absolutisms turn out to be nonexistent (and unlikely to materialize in the future), they would still be worth discussing. The reason is that, in order to evaluate Rawls's theory, we need to assess its overall coherence and plausibility. If the theory is unable to deal adequately with a possible category of states that are peaceful and respect human rights but deny their members a meaningful role in political decision-making, this would still count as a (theoretical) problem worth discussing.³⁷

4. IN OR OUT?

Consider, then, these observations that all seem to be parts of Rawls's theory:

1. Extending the Society of Peoples will help securing lasting peace and stability.
2. Members of the Society of Peoples must
 - a. respect the Law of Peoples, and
 - b. be well ordered.
3. Benevolent absolutisms are neither to be assisted nor sanctioned into either accepting the Law of Peoples (which they might do in any case), or to become well ordered (which they by definition are not).

These observations give rise to some questions. First, it is not clear why well-orderedness is a criterion for inclusion into the Society of Peoples in the first place. This is a general question highlighted by the case of benevolent absolutisms. Well-orderedness, unlike adherence to the Law of Peoples, does not clearly pertain to the overarching goal of peace and stability, since benevolent absolutisms, as noted, are peaceful and most likely law abiding quite generally. Second, it is not clear why benevolent absolutisms, though excluded from the Society of Peoples, are immune to all forms of sanctions, including diplomatic ones. There are no mechanisms, such as assistance or pressure, available to move them closer to inclusion in the Society of Peoples. These states are not to be treated as free and equal parties along with liberal and decent hierarchical societies. In short, benevolent absolutisms appear to occupy some sort of halfway house between the included and the excluded. This makes for a confusing position for these states within Rawls's theory.

This does not amount to any substantial criticism of *The Law of Peoples*. But

37 Something similar, I think, can be said about decent hierarchical societies. Even if no actual state fits the bill, it is important to carve out the limits of international toleration. For a brief discussion of empirical instances of decent societies, see Riker, "The Democratic Peace Is Not Democratic," 620.

it does, I think, give us some reason to search for ways of resolving the uneasy position of benevolent absolutisms. There are at least two natural options. First, one could drop the well-orderedness criterion, and accept as members in good standing of the Society of Peoples all those societies that accept the Law of Peoples, including benevolent absolutisms. Second, one could include the well-orderedness criterion into the Law of Peoples itself, with the foreseeable consequence that benevolent absolutisms cannot adhere to it. This will then provide a clear and sufficient reason not to let them become members of the Society of Peoples. Since it is not obvious (as I argue below) what, if any, gains can be derived from this, this latter option can be further refined by including the crux of the well-orderedness criterion into the set of human rights referred to in principle 6. Specifically, this would entail a *human right* to meaningful political participation. To be sure, meaningful political participation does not comprise the whole of well-orderedness, but this aspect, as shown above, is what benevolent absolutisms lack from becoming well ordered.³⁸ Given the special status Rawls accords to human rights, this would also give members of the Society of Peoples the possibility of subjecting benevolent absolutisms to sanctions, whenever doing so would be reasonable and useful.

In the following, I discuss both these ways of adjusting Rawls's theory, and argue that the latter version of the latter strategy is preferable. Following it would first make *The Law of Peoples* more coherent (both strategies would), and would secure a means of putting pressure on benevolent absolutisms to become well ordered and hence to give their members a meaningful role in political decisions, which would, if successful, in itself represent progress in light of Rawls's stated aim of lasting peace and stability.

5. DROPPING THE WELL-ORDEREDNESS CRITERION

The first strategy, then, would be to simply drop the well-orderedness criterion altogether, and allow benevolent absolutisms as members of the Society of Peoples. If they accept the Law of Peoples, which, as I have argued, it is likely that they would, they should be treated as free and equal parties to the Law of Peoples, and accordingly as members in good standing of the Society of Peoples. Accepting them as equal members, moreover, may serve to secure the goal of continued peace and stability, as both these features are arguably better nurtured within a Society of Peoples than outside it. This will reduce the potential for future resentment and conflict.

38 It is also the only aspect that is not already a part of the Law of Peoples, which requires respect for (the codified) human rights, as well as nonaggression.

There are two problems with this solution, however: one theoretical and one normative. To start with the former, it is unlikely that benevolent absolutisms would be tolerated by liberal societies. Theoretically, this matters, because in Rawls's contractual framework, the Law of Peoples is worked out by liberal and decent societies. If benevolent absolutisms are to be included, they must be accepted by liberal (and decent) peoples as members in good standing. For this to be possible, liberal peoples must be redefined in a way that ensures that they would tolerate states that deny their citizens any meaningful role in political decision-making. Rawls painstakingly argues that liberal societies ought to tolerate decent hierarchical societies. But it appears from his discussion that these societies clearly mark the outer borders of what liberals can reasonably tolerate.³⁹ The issue of toleration does not link directly to the Law of Peoples itself, as this law, as noted, does not contain any requirements concerning the internal structure of participating states apart from the human rights criterion. But it is indirectly connected exactly in the sense that liberals could not reasonably accept the inclusion into the Society of Peoples of any societies that they cannot tolerate. Liberal reluctance to tolerate benevolent absolutisms, then, makes for a tension that would effectively block this strategy.

From a normative point of view, I think there is reason to resist tolerance for benevolent absolutisms. Liberals, regardless of the details of Rawls's theory, should not accept as members in good standing societies that deny their citizens any meaningful role in political decision-making processes. To tolerate benevolent absolutisms would entail tolerating a form of oppression. It seems to me, at least, that denying members of society any political participation amounts to oppression. Further, given the fact that tolerance of decent hierarchical societies is, rightly, controversial, tolerance of benevolent absolutisms seems clearly out of bounds for a liberal theory of international justice.⁴⁰ It appears then, that dropping the well-orderedness criterion is not the way to go.⁴¹

39 Rawls, *The Law of Peoples*, 59–60.

40 See, however, Maffettone, "Should We Tolerate Benevolent Absolutisms?" for an argument to the contrary.

41 As indicated, Rawls has been heavily criticized for his tolerance of *decent* hierarchical peoples. See Tan, "Liberal Toleration in Rawls's Law of Peoples"; Beitz, "Rawls's Law of Peoples"; Pogge, "An Egalitarian Law of Peoples"; Tesson, "The Rawlsian Theory of International Law"; Caney, "Cosmopolitanism and the Law of Peoples"; and Kuper, "Rawlsian Global Justice." For more sympathetic readings, see Bernstein, "A Human Right to Democracy?" and Reidy, "Rawls on International Justice." Some of the criticisms were directed toward the first version of the theory (Rawls, "The Law of Peoples"). In the later version, Rawls elaborates on his arguments for tolerating decent societies. See Huseby, "Liberalism, Tolerance, and Human Rights," for a discussion.

One might reply that decent hierarchical states also engage in a form of oppression, since their members are also denied some crucial rights, including the full right to free speech and the right to democratic participation. This is true, but my argument is primarily negative. I claim that benevolent absolutisms are oppressive to the extent that they should not be tolerated by liberals, and that they should be excluded from the Society of Peoples. This is sufficient to reject the suggestion that they should be included. My argument, however, implicitly presupposes that there is a relevant difference between benevolent absolutisms and decent societies (that is, relevantly different degrees of oppression). In order to rule out the possibility that my claim that benevolent absolutisms should be excluded somehow implies that decent societies should be excluded as well. And it seems that the difference between having meaningful political participation and not having any political participation at all is relevant in the required way. To see this, consider the fact that citizens in systems with political participation have the opportunity to express dissent.⁴² This opportunity is crucial for the possibility of change and reform over the long term. Even though systems with consultation hierarchies fall far short of democratic political participation, they provide a potential for political empowerment that goes significantly beyond systems in which they are lacking.

Surely there could nevertheless be other reasons to exclude decent societies. One could hold that they are *sufficiently* oppressive (though less oppressive than benevolent absolutisms) to be excluded. As noted, many critics have suggested that this is the case. I do not consider this here, however, because my question is whether benevolent absolutisms should be included, not whether decent societies should be excluded.

6. STRENGTHENING THE WELL-ORDEREDNESS CRITERION

An alternative is to alter the Law of Peoples in such a way that benevolent absolutisms can no longer adhere to it. Doing so would render benevolent absolutisms clear non-compliers, and possibly legitimize the use of pressure or sanctions of some kind in order to make them comply. This would be conducive to the aim of improving the conditions of the members of these states.

6.1. Including Well-Orderedness into the Law of Peoples

This could perhaps be achieved by including a demand for well-orderedness into the Law of Peoples, as a principle 9, say. This would effectively block the possibil-

⁴² Rawls, *The Law of Peoples*, 72.

ity for these states to accept the Law of Peoples, and would involve no demand for the liberal toleration of absolutist political systems. However, this would not in and of itself answer the question of how to relate to benevolent absolutisms. As noted, it is not clear whether Rawls presupposes these societies' acceptance of the original Law of Peoples. I have suggested that they might very well accept it, but it is not apparent that this would make any difference with regard to how Rawls thinks they should be dealt with, partly for reasons that have to do with his understanding of human rights (see below).⁴³ If well-orderedness was made a part of the Law of Peoples, benevolent absolutisms would still be excluded from the Society of Peoples, but it is uncertain whether they for that reason could be subject to pressure of any kind, even if they were unable to abide by the (revised) Law of Peoples.

One reason this is uncertain is that violating the Law of Peoples does not automatically open the door to sanctions. A state that violates certain treaties and undertakings (principles 2 and 3, for instance), is not obviously a legitimate target of sanctions, depending of course on the nature and importance of these treaties and undertakings. Violation of other principles, concerning nonintervention and legitimate reasons for warfare, are more likely to trigger sanctions. It is worth repeating here that respecting human rights is sufficient to immunizing a state from sanctions.⁴⁴ Thus, it is not given that the violation of a principle demanding well-ordered institutions would provide grounds for pressuring benevolent absolutisms into compliance.

6.2. *Meaningful Political Participation as a Human Right*

We should consider the possibility of making the crux of the well-orderedness criterion a part of the human rights package referred to in principle 6. In other words, we could argue that there is a human right to meaningful participation in political decision-making.⁴⁵ Rawls's list of human rights is quite limited. It lacks, among other things, a full right to freedom of expression. This and other omissions have, with good reason, been criticized by several theorists.⁴⁶

43 For a related discussion, see Maffettone, *The Coherence and Defensibility of Rawls's Law of Peoples*, 143–44.

44 As noted, it seems clear that external aggression too might trigger sanctions.

45 For relevant discussion, see Peter, "The Human Right to Political Participation."

46 See note 41. Further, several authors argue in favor of a human right to *democracy*, a right which, reasonable though it seems, would represent a much larger alteration to the Rawlsian framework. Christiano, among others, holds that there is a human right to democracy ("An Instrumental Argument for a Human Right to Democracy"). Requiring democratic institutions would mean that decent hierarchical societies could no longer accept the Law of Peoples and would be excluded from the Society of Peoples.

However, some authors suggest that the Law of Peoples *already* entails a human right to political participation.⁴⁷ In my view, this is hard to square with what Rawls says about human rights and benevolent absolutisms. According to Rawls, even though “a benevolent absolutism does respect and honor human rights, it is not a well-ordered society, since it does not give its members a meaningful role in making political decisions.”⁴⁸ This does not seem easily compatible with the view that benevolent absolutisms in fact violate at least one human right—namely the human right to meaningful political participation. In my opinion, Rawls is here best understood on the assumption that there is not a human right to political participation. If there was, this would surely have been worth mentioning explicitly. Notice that while I disagree with Reidy’s interpretation, I obviously agree with the upshot of his view—namely that the Law of Peoples *should* indeed include a right to political participation. I only disagree that this feature is *already* present in Rawls’s theory. Note also that, in one of the three places in which he explicitly mentions benevolent absolutisms, Rawls says that they respect “most” human rights.⁴⁹ I do not think that this is sufficient to establish either that it is the right to political participation that they in particular fail to respect, or that the right to political participation is a human right (on Rawls’s view).

Reidy further suggests that Rawls acknowledges two distinct lists of rights: one minimal list, the respect for which secures the common human good and renders a state immune to sanctions, and a more extensive list, the respect for which suffices for social cooperation and for full toleration in the international society.⁵⁰ This is an interesting suggestion, but I do not share Reidy’s view of the textual basis for this claim. It is true that Rawls lists human rights in two places, and that the lists are not identical, but it seems to me that the second list is a *specification* of the shorter list.⁵¹ That said, Reidy is correct to point out that Rawls’s notion of human rights is more expansive than many critics have realized, but not, I maintain, so expansive as to include the human right to meaningful political participation.

Incorporating the well-orderedness criterion into the list of human rights

47 See, in particular, Reidy, “Human Rights and Liberal Toleration.” Also, Riker suggests that there is a human right to political participation, in “The Democratic Peace Is Not Democratic,” 625–26. However, he seems to later have abandoned this view, in “Human Rights without Political Participation?” 372.

48 Rawls, *The Law of Peoples*, 92.

49 Rawls, *The Law of Peoples*, 63.

50 Reidy, “Human Rights and Liberal Toleration,” 293.

51 Rawls, *The Law of Peoples*, 80n23 and 65.

would, in addition to simply being plausible, accommodate many of the concerns Rawls has in developing his theory. Importantly, it would not affect tolerance of decent hierarchical peoples, nor their adherence to and acceptance of the Law of Peoples. Despite the objections of many critics, Rawls (and some of his defenders) think that tolerating these societies is highly important with regard to the goal of achieving lasting peace. Since decent hierarchical peoples do give their members a meaningful role in political decisions (on Rawls's definition of meaningful), their status would remain unchanged.

Further, this move would be largely in line with the special roles Rawls assigns to human rights:

1. Their fulfillment is a necessary condition of the decency of a society's political institutions and of its legal order.
2. Their fulfillment is sufficient to exclude justified and forceful intervention by other peoples, for example by diplomatic and economic sanctions, or in grave cases by military force.⁵²
3. They set a limit to the pluralism among peoples.⁵³

The inclusion of well-orderedness (still in the limited sense of giving members a meaningful role in making political decisions) into the list of human rights would not touch the first point. The domain of necessary conditions of decency would just expand to some extent.

One problem with the suggestion might be that Rawls explicitly says that human rights is a label reserved for a "special class of urgent rights."⁵⁴ Thus the question is whether a right to political participation is a right with the required level of urgency, which is of course hard to answer conclusively. As I have argued above, however, denying members a right to political participation amounts to a form of oppression, and this seems very urgent indeed. Without such a right, fully competent citizens are coercively deprived of any serious political influence over the societies in which they lead their lives. It is also worth noting that Rawls writes that "laws supported merely by force are grounds for rebellion and resistance."⁵⁵ Such laws are contrasted with laws imposed in a system of political participation. In my view, this indicates that a right to political participation is urgent, even in Rawls's own estimation.

The inclusion would, however, affect the second point in the following crucial

52 Again, this seems true only so far as human rights are concerned. External aggression will also provide grounds for sanctions.

53 Rawls, *The Law of Peoples*, 80.

54 Rawls, *The Law of Peoples*, 79.

55 Rawls, *The Law of Peoples*, 66.

way: any expansion of the list of human rights would provide *additional* grounds for diplomatic, economic, or military intervention. Such expansions should be made with great care. In my view, as noted, holding that members of society should have a human right to a meaningful role in making political decisions in the society in which they lead their lives is a very reasonable claim. It seems hard to oppose it on moral grounds. Also from the perspective of *The Law of Peoples*, this makes sense. The reason is that benevolent absolutisms are not to be tolerated by liberal and decent peoples, precisely because they deny their members such a meaningful role.

More pragmatically, one could ask whether such an expansion of the list of human rights would risk *increasing* international conflict and strife. To the extent that human rights violations may permissibly be met with sanctions, any expansion of human rights provides additional possibilities for conflict. However, this risk must be weighed against the gains, which primarily lie in establishing mechanisms through which to influence benevolent absolutisms into becoming members of the Society of Peoples and adherents to (in my view a more reasonable) Law of Peoples. Further, the question should be whether a candidate human right is plausible or not, and not simply whether such a right may risk provoking conflict.

It is also worth keeping in mind that the use of sanctions is *permitted* on Rawls's account, not obligatory. Whether to impose sanctions should be considered on a case-by-case basis, with reference to whether or not sanctions would be likely to improve the chances of making a benevolent absolutism well ordered. In cases in which the use of sanctions is likely to make things worse, in terms of either the target state's compliance or the conditions for the citizens of the target state, sanctions should not be employed. (Exceptions to this are conceivable, for instance, due to the potential indirect effects of sanctions.⁵⁶) Notice, however, that my aim here is to argue that some mechanism ought to be available, rather than figuring out the conditions under which this mechanism ought to be employed.

Further, since benevolent absolutisms are far less problematic than outlaw regimes, which are externally aggressive, there would be no question of resorting to military intervention, and only rarely to economic sanctions.⁵⁷ A detailed discussion of the justifiability of sanctions lies beyond the scope of the present paper. Note, however, the possibility of smart sanctions that explicitly target regime leaders through such measures as travel bans or asset freezes. Sanctions of this kind may be permissible, at least when they are likely to be both humane

56 Drezner, "The Hidden Hand of Economic Coercion."

57 Thus, benevolent absolutisms' right to self-defense would not be undermined.

and efficient, which will not always be the case.⁵⁸ So the question is really whether the use of relatively mild sanctions would be plausible against states that deny their members any meaningful role in making political decisions. In my view this would indeed be plausible.

It could further be argued that the inclusion of well-orderedness would make the third role of human rights more coherent. I take it that Rawls by “pluralism” refers to “reasonable pluralism,” since human rights obviously do not limit pluralism in any descriptive sense. If this is true, the third role of human rights is better served by the expanded list than the original one. The expanded list is such that those societies that respect the human rights it contains meet the threshold for inclusion into the Society of Peoples, and it is these societies that constitute reasonable pluralism on the international level. The original list is most likely respected by at least one kind of state that falls outside—namely, benevolent absolutisms.

Lastly, and particularly relevant in the present context, the addition would clarify the standing of benevolent absolutisms. They would fall into the category of states that are unwilling to abide by the Law of Peoples, and as a result they are rightfully excluded from being members in good standing of the Society of Peoples. Adherence to the Law of Peoples, moreover, would be the *sole* criterion of membership. In this scenario, it is also the case that liberal and decent societies have the means to pressure benevolent absolutisms into compliance, something that they lack in Rawls’s original account.

6.3. *A Modification or an Abandonment of the Rawlsian Framework?*

One might ask whether the suggested alteration really amounts to an abandonment of *The Law of Peoples*, rather than merely a modification. Consider Rawls’s methodology, according to which, at the international level, peoples enter into a contract that is to protect their interests and regulate their interaction. This methodology is likely to have difficulties coming to grips with states, like benevolent absolutisms, that are externally peaceful but internally oppressive (in the sense of denying political participation rights to its citizens). Since these states pose no threat to the contracting parties, it is not really clear how internal matters, such as human rights, enter the picture in the first place.

The answer, I think, lies again in the special role that Rawls assigns to human rights. These rights have a universal moral force that extends to all societies. While the contracting parties have no *prudential* interest in the internal organization of other societies, they have, by definition, a *moral* interest in the universal

58 For a critical discussion of smart sanctions, see Drezner, “Sanctions Sometimes Smart.”

protection of human rights.⁵⁹ Arguably, my proposal does not fundamentally challenge this picture, but only modifies it, in the sense that minimal political participation rights are included in the basic human rights that set a limit to reasonable pluralism. This seems plausible and in line with the general thrust of Rawls's approach.

7. INTERNAL AND EXTERNAL SELF-DETERMINATION

Maffettone, as noted, has also discussed the theoretical and practical problems posed by benevolent absolutisms. He, too, wonders why these societies are excluded from the Society of Peoples despite the fact that they seemingly would be able and willing to abide by the statutes of the Law of Peoples. His answer starts with the observation that the Law of Peoples supposes externally self-determining (politically autonomous) parties. This is clear from how the statutes are formulated as they refer to free and independent parties, sovereignty, non-intervention, and so on. Next, he argues that Rawls implicitly accepts the idea (found in international law), that external self-determination is premised on *internal* self-determination. Internal self-determination, moreover, requires some measure of collective political participation.⁶⁰ As we have seen, this participation need not be democratic; it can also take the form of, for instance, the consultation hierarchies mentioned earlier.

The lack of internal self-determination, then, explains why benevolent absolutisms are not included in the Society of Peoples. Since they do not have internal self-determination, they cannot have external self-determination. In one sense, this explanation resembles the second alternative discussed above. Even though Maffettone does not propose internal self-determination (understood as limited collective political participation) as an explicit demand in the Law of Peoples, it is clear that this lack of participation is what warrants exclusion. This solution is interesting, but as with the similar solution discussed above, it has little bearing on how liberal states should respond to benevolent absolutisms.⁶¹

Maffettone suggests that liberal and decent societies are permitted to provide benevolent absolutisms with *incentives* in order to encourage them to become liberal or decent, and hence eventually members of the Society of Peoples.⁶²

59 Rawls, *The Law of Peoples*, 27.

60 Maffettone, *The Coherence and Defensibility of Rawls's Law of Peoples*, 146–49.

61 Let me also note that Maffettone does not explain in detail *why* external self-determination should be premised on internal self-determination.

62 Maffettone, *The Coherence and Defensibility of Rawls's Law of Peoples*, 151–52, and “Benevolent Absolutisms, Incentives and Rawls's *The Law of Peoples*.”

Moreover, if the members of the benevolent absolutisms are actively protesting and making demands for more participation, liberal and decent societies can employ “diplomatic pressure mixed with positive incentives.”⁶³ Providing incentives might be useful in many contexts, and there appears to be nothing in *The Law of Peoples* that precludes this. However, it is not immediately clear what “diplomatic pressure,” short of sanctions, might amount to.⁶⁴

All in all, I think that adding the right to political participation to the list of human rights, as I have suggested, is preferable to Maffettone’s proposal. The reason is that my strategy opens a wider range of available tools, including diplomatic and economic sanctions. While these tools will not always work and should not always be employed, it seems better to have a wider rather than a narrower set of options. This solution will make liberal and decent peoples better equipped to strive for the goal of an expanded Society of Peoples.

8. CONCLUDING REMARKS

In this paper, I argued that benevolent absolutisms occupy an uneasy position in Rawls’s *The Law of Peoples*. In an attempt to remedy this unease, I discussed two ways of adjusting the theory. First, I considered including these states into the Society of Peoples. This solution is problematic mainly because it would implausibly strain the notion of liberal tolerance. Second, I considered merging the two criteria for membership in the Society of Peoples—well-orderedness and adherence to the Law of Peoples—by making the latter a part of the former. This in itself does not appear to solve many problems, and I have therefore considered the further move of including a part of the well-orderedness criterion (the right to meaningful political participation) into the very conception of human rights contained in the Law of Peoples. This, I have argued, does clarify the position of benevolent absolutisms. These states are no longer in a situation in which they meet one of two criteria for inclusion in the Society of Peoples. They now fail the only criteria there is. Further, making (this aspect of) well-orderedness a part of the conception of human rights opens the possibility of subjecting benevolent absolutisms to sanctions, since Rawls argues that respecting human rights excludes justified sanctions. All this, of course, rests on whether it is plausible to see the claim to meaningful participation as a human right with universal political and moral force. In my view, this is clearly plausible.

One might wonder, more generally, how the suggested alteration would impact the principles of international law. First, it is clear that Rawls’s account lim-

63 Maffettone, *The Coherence and Defensibility of Rawls’s Law of Peoples*, 153.

64 Note that my account is compatible with the use of incentives, whenever that will be useful.

its external and internal sovereignty as traditionally conceived. This is, as he says, in line with recent trends in international law.⁶⁵ My suggested alteration would be a step further in the sense that states are under a legal obligation to allow their citizens a meaningful role in political decision-making.⁶⁶ Moreover, given that Rawls ties human rights violations and sanctions so tightly together, any expansion of human rights will automatically increase the grounds for diplomatic, economic, and military intervention. In international law, the principle of non-intervention is not set in stone, but it seems that allowing intervention in order to pressure states into giving their citizens a meaningful role in political decision processes would most likely imply a further limitation of state sovereignty.⁶⁷ As I already noted, such a limitation has to be considered and implemented very carefully.⁶⁸

University of Oslo
roberthu@stv.uio.no

REFERENCES

- Beitz, Charles R. "Rawls's Law of Peoples." *Ethics* 110, no. 4 (July 2000): 669–96.
- Bernstein, Alyssa R. "A Human Right to Democracy? Legitimacy and Intervention." In *Rawls's Law of Peoples: A Realistic Utopia?* edited by Rex Martin and David A. Reidy, 278–99. Oxford: Blackwell, 2006.
- Caney, Simon. "Cosmopolitanism and the Law of Peoples." *Journal of Political Philosophy* 10, no. 1 (March 2002): 95–123.
- Christiano, Thomas. "An Instrumental Argument for a Human Right to Democracy." *Philosophy and Public Affairs* 39, no. 2 (Spring 2011): 142–76.
- Drezner, Daniel W. "The Hidden Hand of Economic Coercion." *International Organization* 57, no. 3 (Summer 2003): 643–59.

65 Rawls, *The Law of Peoples*, 27.

66 Note that some human rights documents make reference to the right of political participation (Reidy, "Human Rights and Liberal Toleration").

67 In this paper, I take much of Rawls's framework for granted. In a subsequent discussion it could be interesting to ask whether the connection between human rights and sanctions ought to be as closely connected as Rawls assumes.

68 I would like to thank Kim Angell, Ludvig Beckman, Lene Bomann-Larsen, Göran Duus-Otterström, Eli Feiring, Fredrik Dybfest Hjorthen, Andreas Hvidsten, Jakob Elster, Aaron Maltais, Grethe Netland, Pietro Maffettone, Håvard Strand, and two anonymous referees for very helpful comments on earlier drafts of this paper.

- . “Sanctions Sometimes Smart: Targeted Sanctions in Theory and Practice.” *International Studies Review* 13, no. 1 (March 2011): 96–108.
- Estlund, David. *Democratic Authority*. Princeton: Princeton University Press, 2008.
- Freedom House. *Freedom in the World 2016*. https://freedomhouse.org/sites/default/files/FH_FITW_Report_2016.pdf.
- Huseby, Robert. “John Rawls and Climate Justice: An Amendment to *The Law of Peoples*.” *Environmental Ethics* 35, no. 2 (Summer 2013): 327–44.
- . “Liberalism, Tolerance, and Human Rights.” *Nordic Journal of Human Rights* 25, no. 3 (2007): 245–58.
- Kuper, Andrew. “Rawlsian Global Justice: Beyond the Law of Peoples to a Cosmopolitan Law of Persons.” *Political Theory* 28, no. 5 (October 2000): 640–74.
- Maffettone, Pietro. “Benevolent Absolutisms, Incentives, and Rawls’s *The Law of Peoples*.” *Politics, Philosophy and Economics* 15, no. 4 (November 2016): 379–404.
- . *The Coherence and Defensibility of Rawls’s Law of Peoples*. PhD Diss., London School of Economics, 2013.
- . “Should We Tolerate Benevolent Absolutisms?” *Social Theory and Practice* 42, no. 3 (July 2016): 535–54.
- Marshall, Monty G., Ted Robert Gurr, and Keith Jagers. *The Polity IV Project: Political Regime Characteristics and Transitions, 1800–2016*. Center for Systemic Peace, 2017. <http://www.systemicpeace.org/inscr/p4manualv2016.pdf>.
- Neufeld, Blain. “Liberal Foreign Policy and the Ideal of Fair Social Cooperation.” *Journal of Social Philosophy* 44, no. 3 (September 2013): 291–308.
- Peter, Fabienne. “The Human Right to Political Participation.” *Journal of Ethics and Social Philosophy* 7, no. 2 (February 2013): 1–16.
- Pettit, Philip. “Rawls’s Peoples.” In *Rawls’s Law of Peoples: A Realistic Utopia?* edited by Rex Martin and David A. Reidy, 38–55. Oxford: Blackwell, 2006.
- Pogge, Thomas W. “An Egalitarian Law of Peoples.” *Philosophy and Public Affairs* 23, no. 3 (July 1994): 195–224.
- Rawls, John. *Justice as Fairness: A Restatement*, edited by Erin Kelly. Cambridge, MA: Harvard University Press, 2001.
- . “The Law of Peoples.” In *Human Rights: Oxford Amnesty Lecture*, edited by Stephen Shute and Susan Hurley, 41–82. New York: Basic Books, 1993.
- . *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.
- . *Political Liberalism*. New York: Columbia University Press, 1993.
- . *A Theory of Justice*. Oxford: Oxford University Press, 1971.
- Reidy, David A. “Human Rights and Liberal Toleration.” *Canadian Journal of Law and Jurisprudence* 23, no. 2 (July 2010): 287–317.

- . “Rawls on International Justice: A Defense.” *Political Theory* 32, no. 3 (June 2004): 291–319.
- Riker, Walter J. “The Democratic Peace Is Not Democratic: On Behalf of Rawls’s Decent Societies.” *Political Studies* 57, no. 3 (October 2009): 617–38.
- . “Human Rights without Political Participation?” *Human Rights Review* 15, no. 4 (December 2014): 369–90.
- Soper, Philip. *A Theory of Law*. Cambridge, MA: Harvard University Press, 1984.
- Tan, Kok-Chor. “Liberal Toleration in Rawls’s Law of Peoples.” *Ethics* 108, no. 2 (January 1998): 276–95.
- Téson, Fernando R. “The Rawlsian Theory of International Law.” *Ethics and International Affairs* 9, (March 1995): 79–99.
- Wood, Reed M., and Mark Gibney. “The Political Terror Scale (PTS): A Re-Introduction and a Comparison to CIRI.” *Human Rights Quarterly* 32, no. 2 (May 2010): 367–400.

SHARED INTENTION IS NOT JOINT COMMITMENT

Matthew Kopec and Seumas Miller

MARGARET GILBERT has long defended the view that, roughly speaking, agents share the intention to perform an action if and only if they jointly commit to performing that action.¹ This view has proven both influential and controversial. While some authors have raised concerns over the joint commitment view of shared intention, including at times offering purported counterexamples to certain aspects of the view, straightforward counterexamples to the view as a whole have yet to appear in the literature.² Here we provide such counterexamples to show that joint commitment is neither necessary nor sufficient for shared intention.

Before presenting the counterexamples, we must explain Gilbert's joint commitment view of shared intention. As she states:

Members of some population *share an intention* to do A if and only if they are jointly committed to intend as a body to do A.³

The key notion here is that of joint commitment.⁴ As Bratman points out, Gilbert sees joint commitment as a somewhat basic notion, and it would therefore

- 1 Gilbert, *Joint Commitment*, "Shared Intention and Personal Intentions," *Sociality and Responsibility*, "Walking Together," and "What Is It for Us to Intend?"
- 2 For general critiques of the joint commitment view of shared intention, see, e.g., Bratman, *Shared Agency*, ch. 5; Ludwig, *From Individual to Plural Agency*, 261–71; and Miller, review of *Joint Commitment*. For purported counterexamples, see, e.g., Bratman, "Shared Intention," 110–11; *Faces of Intention*, 132–33; "Dynamics of Sociality," 7; "Modest Sociality and the Distinctiveness of Intention," 152–53; and *Shared Agency*, 112, 116. Also see Ludwig, *From Individual to Plural Agency*, 266–68. We later discuss how Gilbert has some reasonable routes to reply to some of these kinds of examples that she lacks with our counterexamples. We thank an anonymous referee for leads to some of these examples, which we had missed in an earlier draft.
- 3 Gilbert, "Shared Intention and Personal Intentions," 179.
- 4 What it could mean to "intend as a body" also calls out for further explanation, but we skip this in the interest of space, since it will not matter to the counterexamples we present. See Gilbert (*Joint Commitment*, 115–17) for discussion of this aspect of joint commitment.

be pointless to attempt to give a full and proper definition of the concept.⁵ But Gilbert does give her readers some guidance to fix upon the notion. For example, she states, “Intuitively an appropriate *agreement* between the parties is sufficient to bring a shared intention into being.”⁶ It follows from the biconditional above that an appropriate agreement between the parties will also be sufficient to bring about a joint commitment. What makes an agreement *appropriate*? She states that “each of two or more people must openly express his personal readiness jointly with the others to commit them all in a certain way . . . Once the concordant expressions of all have occurred and are common knowledge between the parties, the joint commitment is in place.”⁷ This suggests that Gilbert holds the following:

Agreement: If all relevant parties openly agree to perform a collective action *A* by expressing their readiness to each other to commit to acting in the relevant ways, then the parties are jointly committed to doing *A*.⁸

Thus, the right kind of agreement is a sufficient condition for a joint commitment on Gilbert’s account.

Gilbert offers some further guidance on the notion of joint commitment by giving what are, and are not, some necessary conditions for a joint commitment to obtain. We start with the latter. Gilbert accepts what she calls the “disjunction criterion,” which states that, “when two or more people share an intention, none of them need have a personal contributory intention.”⁹ She supports this claim with her classic example of Ned and Olive, who agree to walk together to the top of a hill.¹⁰ Gilbert argues that even if both of them change their minds along the way, thus losing their personal intentions to summit the hill, the pair can continue to have the shared intention to summit the hill. Since, according to Gilbert, individual personal intentions to perform the action are not a necessary condition for them to hold a shared intention, Gilbert’s biconditional above further entails that such personal intentions are not necessary for a joint commitment. Thus, she holds the following:

Disjunction: A group can be jointly committed to perform the collective

5 Bratman, *Shared Agency*, 113–14.

6 Gilbert, “Shared Intention and Personal Intentions,” 169.

7 Gilbert, “Shared Intention and Personal Intentions,” 180.

8 To be fully comprehensive, a common knowledge condition must also be added here, but this will not be relevant to anything that follows, so we pass over it in the interest of space.

9 Gilbert, “Shared Intention and Personal Intentions,” 171.

10 Gilbert, “Shared Intention and Personal Intentions,” 171–72.

action *A* even though no members of the group personally intend to act in order to ensure that *A* obtains.

As far as what is necessary for a joint commitment to obtain, Gilbert focuses on the obligations that are always carried with joint commitments. She states that “an adequate account of shared intention will entail that each party to a shared intention is obligated to each to act as appropriate to the shared intention in conjunction with the rest.”¹¹ Gilbert argues that her joint commitment account of shared intention meets this criterion, by presenting cases in which a joint commitment to do some act is in place, yet one of the members fails to do her part in performing the action. Given that it seems fully justified for the other parties to the joint commitment to rebuke the bad actor, this suggests that joint commitments give rise to the relevant obligations. But it is crucial to Gilbert’s view that the kind of obligation involved is not of the *moral* or *institutional* variety.¹² This suggests that Gilbert accepts the following:

Rebuke: A group is jointly committed to perform a collective action *A* only if members of the group are justifiably subject to (non-moral, non-institutional) rebuke for failing to do their part in performing the action.

Thus, the appropriateness of rebuke (of a certain sort) upon failing to do one’s part is a necessary condition for joint commitment.

So, while Gilbert does not provide her reader with a full definition of joint commitment, she does give enough details to fix upon what she intends by the phrase. We now show that shared intention is not joint commitment.

1. JOINT COMMITMENT IS NOT SUFFICIENT FOR SHARED INTENTION

In this section, we show that joint commitment is not sufficient for shared intention. We proceed by simply giving a counterexample. Take the following case:

Saboteurs: A government during wartime asks for volunteers to build a strategically important bridge. Unbeknownst to the government, every eventual volunteer in the group turns out to be a conscientious objector; each privately intends to thwart the government’s efforts to win the war whenever possible. That said, every member publicly assents to the goal of building the bridge as a team, agreeing to each other that they each will do their part. Moreover, their outward behavior looks as though the team is working as a well-functioning body. But each privately intends to

11 Gilbert, “Shared Intention and Personal Intentions,” 171.

12 See Miller, review of *Joint Commitment*.

merely act as though she is helping to build the bridge. Each occasionally commits various blunders, losing a crucial piece of equipment here, incorrectly mixing some concrete there, etc. The result is that the bridge is never completed.

Now, suppose the government liaison, i.e., the official overseeing how the project is going, realizes what is going on.¹³ If her irate superior officer calls for an explanation of the bridge's delay, we could imagine her retorting, "The team of volunteers you sent me don't really share any intention to build this bridge." We find it obvious that what the project leader says here is true. The dissidents do not share the intention to build the bridge; rather they each personally have the intention to merely *seem* like they are building the bridge.¹⁴

While it is clear that the group of conscientious objectors lacks a shared intention to build the bridge, Gilbert would be forced to say they do have a joint commitment to build the bridge. Recall that Gilbert accepts *Agreement*, and the volunteers in this case have publicly expressed their readiness to each other to commit to acting in the relevant ways so that they build the bridge together. Since such an agreement is sufficient for a joint commitment on Gilbert's account, Gilbert must say that the volunteers are jointly committed to building the bridge. And since the builders are jointly committed to building the bridge, but obviously do not share the intention to build the bridge, it follows that joint commitment cannot be a sufficient condition for shared intention.

Gilbert might object that, in the *Saboteurs* case, publicly assenting to the goal of building the bridge together is not enough to constitute a joint commitment, perhaps because an individual cannot jointly commit to a goal unless she has some intention to actually attain the goal. Unfortunately, this line is not open to Gilbert, since she accepts *Disjunction*. Recall that, on her account, one cannot determine whether a group has a joint commitment by examining whether the

13 We are assuming that the government liaison is not one of the volunteers.

14 An anonymous referee has pressed us to explain the difference between our example and one presented by Bratman (*Shared Agency*, 112, 116) involving a pair who insincerely promise to plow the commons together. First, Bratman does not intend his case as a general counterexample to Gilbert's view that shared intention is joint commitment. Rather, he uses it to argue that potential joint actors can have obligations to act jointly without sharing the intention to act jointly. As he puts it, "Since we have each promised—albeit insincerely—we each have a moral obligation to the other to plow" (112). Second, in our case there is no promising, and thus, we would argue, no corresponding moral obligations. Thus, Gilbert has a response to Bratman's case that she lacks with ours. Since the kinds of obligations at issue in joint commitment are not *moral* obligations, she could simply claim that his case is one in which moral obligations are present but a joint commitment is not. We will not take a stand on whether such a response is plausible. We thank an anonymous referee for pressing us here.

agents have a personal intention to perform the action. While Gilbert seems to think that *Disjunction* is a positive feature of her view, *Saboteurs* shows that in fact this requirement introduces a flaw.

But there might be a more nuanced way for Gilbert to press this kind of objection. In the Ned and Olive case, which Gilbert uses to support *Disjunction*, both Ned and Olive do *initially* intend to climb the hill, even though their personal intentions fade at some point during the climb. Gilbert explains that when she says the members must be personally ready to commit to acting she means to imply “that each is indeed personally ready for this, and that he expresses this readiness.”¹⁵ Perhaps when Gilbert talks of an individual being “personally ready” to commit, she means that the individual must, at the point of expressing the commitment, actually have a personal intention to perform the relevant acts. Since the conscientious objectors in the *Saboteurs* case do not intend to build the bridge from the very start, Gilbert might be able to use this move to claim that there cannot have been a genuine joint commitment. And if there was not a genuine joint commitment, that case cannot serve as a counterexample.

There are two ways to reply to this nuanced version of the objection. The first reply would be to slightly modify the *Saboteurs* case. We could stipulate that the volunteers really did intend to build the bridge at the outset, but later found out that their government had embarked on an unjust war, and only at this later time became conscientious objectors and commenced to sabotage the project. We still feel it would be obviously true when the liaison reports, “The team of volunteers you sent me don’t really share any intention to build this bridge.” A second reply would point out that the revised objection carries a heavy cost for Gilbert. What is distinctive about Gilbert’s view is that she is able to give an account of shared intention that is not held captive to personal intentions, which she sees as a benefit for her view over the views of her competitors. On the revision, her account of shared intention is held captive to personal intentions, although only at the point when the joint commitment is made. This move, while it might help get some of the intuitions right, seems to us to be otherwise unmotivated. It is hard to see why personal intentions could be so very important in setting up the shared intention and yet so very unimportant in maintaining it. This suggests to us that views like those of Bratman and Miller, which link shared intentions to personal intentions *for the duration*, are preferable to Gilbert’s under this revised reading.¹⁶

15 Gilbert, *Joint Commitment*, 115.

16 Bratman, “Shared Intention” and *Shared Agency*; Miller, “Joint Action” and *Social Action*.

2. JOINT COMMITMENT IS NOT NECESSARY FOR SHARED INTENTION

In this section, we show that joint commitment is not necessary for shared intention. Before we present a counterexample, we need to make a few preliminary remarks. All of the players in the debate over the nature of shared intention agree that shared intention is a precondition of joint action.¹⁷ For example, Gilbert's own account of shared intention was drawn up in order to account for joint actions like walking together.¹⁸ So, it is safe to assume that every joint action will be accompanied by a shared intention.

To see why joint commitment cannot be a necessary condition, consider the following case:

Shopping Spree: Two strangers walk into a grocery store and are immediately stopped by the manager who has randomly chosen them for a store promotion. She states that if the two of them can retrieve all the ingredients for an ice cream sundae and deliver them to checkout number five within twenty seconds of a starting buzzer, they will each receive a five-hundred-dollar store gift certificate. She explains that the necessary ingredients are crushed peanuts, maraschino cherries, whipped cream, and vanilla ice cream, and she then points out large signs marking the aisles where those items are located. Since the cherries and peanuts are on one side of the store and the ice cream and whipped cream are on the other, and given the very short time to complete the task, it is instantly clear to each contestant that they will only be able to complete the task if they each sprint to separate sides of the store to grab those respective items. Before they can discuss any plans, the buzzer sounds, and they each start off to the side of the store they are closest to, quickly looking back to check to make sure the other is doing the same. They successfully return all four items.

Now, clearly these two strangers have successfully completed the joint action

17 See, e.g., Bratman, "Shared Intention" and *Shared Agency*; Miller, "Joint Action" and "Intentions, Ends and Joint Action"; Tuomela, *Social Ontology*; Tuomela and Miller, "We-Intentions." At least, every player in the debate that we are aware of accepts this. To be clear, the phrase "joint action" is used as a term of art in the literature, to be distinguished from mere "collective action." For example, two agents walking together is a joint action, while 380 million Americans creating 4.8 billion metric tons of CO₂ in a year is a mere collective action. Nobody believes the latter entails the existence of a shared intention, for obvious reasons. Shared intention is really what makes joint action distinct.

18 See Gilbert, "Walking Together," "What Is It for Us to Intend?" and "Shared Intention and Personal Intentions."

of delivering the four necessary ingredients to checkout number five within the time limit. And, as we suggested earlier, the fact that this was a joint action entails that the pair had the shared intention to perform the action in question. So, these two strangers each shared the intention to deliver the four necessary ingredients to checkout number five within the time limit.¹⁹

But do the shoppers have a joint commitment to perform the action in question? Because Gilbert accepts *Rebuke*, they have such a commitment only if they each justifiably could rebuke the other for failing to do their part. But this kind of susceptibility to rebuke seems entirely missing in this case. For example, the strangers did not have time to make any commitments to each other. As soon as the starting buzzer rang, they did not stop to publicly assent to each other their goals of doing thus and so. And, being complete strangers, they could not have any standing commitments to acting cooperatively in cases like this. Thus, if one of the shoppers initially started off for the sundae items but then abandoned her plan to participate and simply moved along to collecting the items on her usual shopping list, the other shopper would not have any real standing to rebuke her for not helping to complete the task. Since neither participant in this case has

19 An anonymous referee has pressed us to explain the difference between our example and some presented by Bratman. In a number of places, Bratman uses cases in an attempt to establish that shared intentions need not generate obligations, e.g., (1) his “Ayn Rand” singers who explicitly waive their obligations (“Shared Intention,” 111; *Faces of Intention*, 133; “Modest Sociality and the Distinctiveness of Intention,” 151; see also Ludwig, *From Individual to Plural Agency*, 267–68); (2) people coerced into agreeing to perform joint actions (*Faces of Intention*, 132–33; “Modest Sociality and the Distinctiveness of Intention,” 152); (3) those who share the intention to perform immoral joint actions (“Modest Sociality and the Distinctiveness of Intention,” 152); and (4) cases of casual joint actions like spontaneously applauding a performance (“Dynamics of Sociality,” 7; “Modest Sociality and the Distinctiveness of Intention,” 151; see also Ludwig, *From Individual to Plural Agency*, 266–68). First, as with note 14, above, these examples are not intended as general counterexamples to Gilbert’s view but are rather used for the narrower purpose of showing that shared intentions do not always generate obligations. And although we employ a closely related strategy here, our example is not subject to Gilbert’s readily available replies to these other examples. *Contra* (1)–(3), Gilbert could plausibly claim that while the agents in those cases lack *all-things-considered* obligations to perform the joint actions, due to extenuating circumstances like side agreements or moral considerations, they can nonetheless have *pro tanto* obligations grounded in their agreements (cf. Gilbert, *Joint Commitment*, 112–13). And Gilbert could plausibly argue that the kinds of casual joint actions offered by Bratman and Ludwig in (4) are not genuine instances of joint intentional action. For example, she could plausibly claim that applauding together is more akin to eating a pile of crawfish together than it is to singing a duet together. And most of Ludwig’s examples intuitively lack persistent obligations to act, simply because those examples lack success conditions. But Gilbert’s likely moves to dispute these cases will not work against our example. We thank an anonymous referee for pressing us here.

any obligation to do their part to complete the joint action, Gilbert would have to accept that they do not have any joint commitment to complete the action in this case. It follows that joint commitment cannot be necessary for shared intention.

Gilbert might object that in this case there actually *is* a joint commitment between the strangers by simply denying that the two shoppers lack the obligation to do their part to complete the task. That is, she could claim that, if one of the shoppers were to run off to start her usual shopping as opposed to working to complete the joint action, then the other shopper would indeed be justified in rebuking her. While it is difficult to argue with someone who simply does not share our intuitions, we think it is possible to account for such intuitions while still denying that the shoppers have a joint commitment to perform the action. The first response would be to point out that on Gilbert's account the rebuke must be not only *understandable*, but also *justified*. While the one shopper who does his part might feel a bit irritated that the other decided not to participate, this does not entail that his rebuke would be justified. For example, if the shoppers were enemies as opposed to merely strangers, it becomes a bit more clear that the two do not have any obligation to ensure that a good result is bestowed upon the other. If she does not want the gift certificate herself, nothing binds her to participating. After all, she never agreed to participate.

But a second reply would point out that, as often happens in cases of joint action, various other moral concerns can enter the picture and cloud one's intuition. If it seems intuitively plausible that rebuke would be justified in the case of one failing to do one's part, this might be because the reader takes the shopper to have a standing obligation to improve the welfare of others if the cost to herself is trivial.²⁰ We could probe whether this factor is what is generating the intuition by changing the case so that the "prize" is something that is unlikely to enhance either stranger's welfare, such as that each will be congratulated over the loudspeaker, or some such. In such a case, it should become obvious that a rebuke of the shopper who declines to participate would be completely unjustified. So, if the reader finds it intuitively plausible that each shopper in the original case does have an obligation to do their part, this intuition traces back to a *moral* norm, as opposed to a norm generated directly from a joint commitment.²¹ Therefore, this kind of response on the part of Gilbert is a nonstarter.

20 Similarly, one might think that falsely encouraging someone else to rely on one's behavior when something morally significant is at stake might be morally wrong (cf. Scanlon, "Promises and Practices"), and five hundred dollars' worth of groceries could certainly be morally significant. But, as we point out, moral obligations are dialectically irrelevant here.

21 We could say a similar kind of thing with regard to practical normativity. The one shopper's

3. CONCLUSION

In this essay, we have shown that joint commitment is neither necessary nor sufficient for shared intention. What went wrong? We suggest that in Gilbert's attempt to sever the tie between shared intention and personal intention, which is what her notion of joint commitment is supposed to do, she lost contact with what shared intention truly is. Our counterexamples suggest, somewhat indirectly, that shared intentions and personal intentions are closely tied together. In *Saboteurs*, the group lacks the shared intention to build the bridge because none of them really has the personal intention to help build the bridge. In *Shopping Spree*, the group has the shared intention to deliver the relevant items to the checkout lane in time because each has the personal intention to do their part to complete that joint action (along with meeting other conditions that will not concern us here). This all suggests that accounts of shared intention that are closely tied to personal intentions, such as Bratman's or Miller's, are preferable to Gilbert's joint commitment account.²²

Australian National University
matthewckopec@gmail.com

Charles Sturt University
Delft University of Technology
Oxford University
semiller@csu.edu.au

irritation would be more explicable if we assume that the other shopper actually values the gift certificate than it would be if we assume that she does not value it. This suggests that another way we can make sense of the shopper's irritation is by noting that the other shopper is acting irrationally, i.e., she is not effectively pursuing *her own interests*. But notice that this kind of rebuke is not based on what the one shopper owes the other shopper, but rather what she owes herself. That will not be the kind of rebuke Gilbert needs. We thank an anonymous referee on this point.

- 22 Bratman, "Shared Cooperative Activity," "Shared Intention," and *Shared Agency*; Miller, "Joint Action" and *Social Action*.

We would like to thank Anne Schwenkenbecher, Baskarjit Neog, Luke Roelofs, Mary Walker, and the audience at the 2017 Eastern Division Meeting of the American Philosophical Association for their helpful comments, questions, and discussion. We would also like to thank two anonymous referees for critical comments or suggestions that helped us improve the final paper. Apologies if we have accidentally omitted anyone who deserves recognition. Kopeć's work on this project was supported under an ARC DECRA grant, DE180101119.

REFERENCES

- Bratman, Michael. "Dynamics of Sociality." *Midwest Studies in Philosophy* 30, no. 1 (September 2006): 1–15.
- . *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press, 1999.
- . "Modest Sociality and the Distinctiveness of Intention." *Philosophical Studies* 144, no. 1 (May 2009): 149–65.
- . *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press, 2014.
- . "Shared Cooperative Activity." *Philosophical Review* 101, no. 2 (April 1992): 327–41.
- . "Shared Intention." *Ethics* 104, no. 1 (October 1993): 97–113.
- Gilbert, Margaret. *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press, 2013.
- . *On Social Facts*. Princeton: Princeton University Press, 1989.
- . "Shared Intention and Personal Intentions." *Philosophical Studies* 144, no. 1 (May 2009): 167–87.
- . *Sociality and Responsibility: New Essays in Plural Subject Theory*. Lanham, MD: Rowman & Littlefield Publishers, 2000.
- . "Walking Together: A Paradigmatic Social Phenomenon." *Midwest Studies in Philosophy* 15, no. 1 (September 1990): 1–14.
- . "What Is It for Us to Intend?" In *Sociality and Responsibility: New Essays in Plural Subject Theory*, 65–86.
- Ludwig, Kirk. *From Individual to Plural Agency: Collective Action I*. Oxford: Oxford University Press, 2016.
- Miller, Seumas. "Intentions, Ends and Joint Action." *Philosophical Papers* 24, no. 1 (1995): 51–66.
- . "Joint Action." *Philosophical Papers* 21, no. 3 (1992): 275–97.
- . Review of *Joint Commitment: How We Make the Social World*, by Margaret Gilbert. *Notre Dame Philosophical Reviews*, August 14, 2014. <http://ndpr.nd.edu/news/joint-commitment-how-we-make-the-social-world>.
- . *Social Action: A Teleological Account*. Cambridge: Cambridge University Press, 2001.
- Scanlon, Thomas. "Promises and Practices." *Philosophy and Public Affairs* 19, no. 3 (Summer 1990): 199–226.
- Tuomela, Raimo. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press, 2013.

Tuomela, Raimo, and Kaarlo Miller. "We-Intentions." *Philosophical Studies* 53, no. 3 (May 1988): 367–89.