JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXX · NUMBER 6
October 2025

ARTICLES

OPEN ACCESS FREE FOR AUTHORS AND READERS ISSN 1559-3061

JOURNAL *of* ETHICS
& SOCIAL PHILOSOPHY

VOLUME XXX · NUMBER 6
October 2025

ARTICLES

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY
http://www.jesp.org
ISSN 1559-3061

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

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

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

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

Journal of Ethics and Social Philosophy
vol. 30, no. 6 · october 2025
https://doi.org/10.26556/jesp.v30i6.4417

PARENTS AND PEERS

A KANTIAN MORAL DEVELOPMENT

William Grant Ray

> The man who believes himself endowed with an autonomous will thus places himself in another order of things and relates himself to determining grounds of an entirely different sort from when he perceives himself as a phenomenon. . . . The fact that he has to represent and think everything in this twofold way is not at all contradictory, for it rests in the first place on his consciousness of himself as an object affected by the senses, in the second on the consciousness of himself as intelligence, that is, as an active subject who, in using reason, is freed from any passive attachment to sensory impressions. Thus, in line with the basic orientation of the transcendental method, here too determination of the object is the result of the mediation of the *analysis of judgment*.
>
> —Ernst Cassirer, *Kant's Life and Thought*

How do we become moral persons—performing acts that are our own, endowed with duties, and standing in relations of reciprocal accountability? This is a developmental question. Any satisfying answer will have to capture some change in kind, in a mess of actual changes of apparent degree. Eventually, a number of grains of sand becomes a heap—but it will not be our task to identify that number, nor need we suppose that it is identifiable. Instead, in offering here a narrative of this development, I will discuss the types of experiences that contribute to and render possible the change in kind—that is, those that lead us to our state of dutiful moral responsibility. Specifically, this essay considers two forms of social relations as relevant contributors to development: relations between the child and their authorities (represented by the figure of the parent) and relations between the child and their equals (represented by the figure of the peer). The project is a partisan one: I intend to propose a theory of moral development for a Kantian practical philosophy, but the extent to which this theory is compatible with other approaches to practical philosophy is left to the judgment of the non-Kantian reader.

Twenty-five years ago, David Velleman attempted a project of this kind. In it, he demonstrated the remarkable potential of the Freudian theory of superego formation to explain how we develop into moral persons, in the Kantian sense.[1] Some reasons for this potential present themselves rather readily: a faculty psychology, an identity between that which is ideal and that which legislates and prescribes, the interplay of internal and external authority. But something significant also stood in the way: an ostensibly critical incompatibility of Kantian "rationalistic" commitments, and Freud's "anti-rationalism." Velleman's reading, however, "purged" Freud of his anti-rationalism and led to the identification of a *rational superego*, a kind of personified ideal of practical reason.[2] His innovation consisted in proposing that the superego, an "introjected" figure of parental authority that presents itself as the voice of conscience, could be seen as containing legislative humanity in us.[3] Other work in the Kantian tradition tries, along similar lines, to assure compatibility of Freud and Kant in the realm of moral psychology.[4]

While I think there is deep insight in this Freudian connection, and I think that the idea of a rational superego plays an indispensable role in offering a story of moral development for Kantian practical philosophy, I think it cannot be the *whole story*. The trouble is, the Freudian theory of the superego and its formation is a theory that concerns itself only with an authority relation—the parent-child relation. Kantian practical philosophy ultimately concerns itself with persons standing in relations of moral equality and reciprocal recognition. Principles of equality and reciprocity are not those characterizing relations of authority; the parent is not, in the critical development period, an equal. *Paternalism*, and its discontents, will resultingly be seen as the reason we cannot account for certain necessary features of a moral psychology capable of *autonomous* action, in relations of authority alone. Though we learn from the parents how to subject impulsiveness to rule (a necessary condition of moral action), we do not yet know how to subject impulsiveness to rule recognized as *our own* (a sufficient condition of moral action).

There is, however, another great tradition of thinking through the psychology of moral development that has been neglected in these investigations—one

1 Velleman, "A Rational Superego." Before Velleman, this was noted especially by Samuel Scheffler and John Deigh. More specifically, they thought, rationalist moralities have trouble explaining moral motivation: e.g., "that an authoritative aspect of the self may play a role in moral motivation is not obviously incompatible in itself with the rationalist position. . . . On this view, the superegos of rational human agents confer motivational authority on moral principles in recognition of their status as principles of pure practical reason" (Scheffler, *Human Morality*, 96–97n). See also Deigh, *The Sources of Moral Agency*, ch. 6.

2 Velleman, "A Rational Superego," 558.

3 Introjection is the process by which a child internalizes representational images of others.

4 See especially Longuenesse, *I, Me, Mine*.

that offers us the resources to overcome this paternalism problem—that of the Piagetian stage theory tradition.[5] Equality and reciprocity characterize peer interaction, which models the active construction and justification of norms (rather than passive internalization, as in relations with authorities). In relations with equals, I claim, we have the resources to make sense of a moral psychology properly fit for persons, recognized legislators of the laws of their own acts. But this too cannot be the whole story. Piaget's account conversely denies contributions from the parental relation that we understand as essential to development. The *logic* of the Piagetian story, I claim, is such that even the most rational of all moral notions can emerge in peer interaction in complete autonomy from relations of authority. *Pace* Piaget, this essay will hold that the basis of peer interaction will necessarily include our having gained certain positive psychological resources from authority relations.

We begin in the first section by identifying what features a Kantian theory of development will need to explain and what the developed state of personhood must minimally consist in. We find that it will explain (1) the original partitioning of two aspects or parts of our person (the *legislative* and the *obedient*) and (2)

5 These two traditions, Freudian psychoanalysis and Piagetian stage theory, have been historically the most generative in the empirical psychology of moral development. Conceived sufficiently abstractly, their frameworks encompass a great amount of work. See especially Walker et al., "Parent and Peer Contexts for Children's Moral Reasoning Development." Their object was to compare parent and peer contributions to developing moral maturity, advancing that both are essential to development and "challenging the polarized view that characterizes the field" (1033). They further "hypothesized that the specific processes underlying development would operate differently in these contexts" (1033). For the parents, "inductive discipline, authoritative parenting, responsiveness, and involvement are associated with children's moral maturity" (1033–34). Cf. Hoffman and Saltzstein, "Parent Discipline and the Child's Moral Development"; Hart, "A Longitudinal Study of Adolescents' Socialization and Identification as Predictors of Adult Moral Judgment Development"; Dunton, "Parental Practices Associated with Their Children's Moral Reasoning Development"; and Boyes and Allen, "Styles of Parent-Child Interaction and Moral Reasoning in Adolescence." For the peers, "Kruger and Tomasello ["Transactive Discussions with Peers and Adults"] . . . found peer discussions more transactive [for moral reasoning] than mother/child ones [and Kruger ("The Effect of Peer and Adult-Child Transductive Discussions on Moral Reasoning") found the same]"; the findings "support the cognitive developmental view regarding the significance of peers and provide the rationale for the hypothesis that the challenging, operational type of interactions between peers will foster moral maturity" (1034). Walker et al.'s hypothesis that "both socialization contexts are important for development . . . [and that] given the profoundly different nature of these relationships on a variety of dimensions, the processes that impact on moral development operate differently in stimulating growth" was in the end supported (1044). Our theory prescribes different forms of contributions from each type of relationship, and I offer albeit very limited substantiation in footnotes that some empirical findings evidence the prescribed forms.

how we come to regard ourselves and others as a *unity* of these two parts—that is, how we come to regard ourselves and others as autonomous (*self-legislative*), how we are awakened to our own personhood and the personhood of others.[6] We turn in the second section to discuss the parental relation through critical engagement with a Freudian account, represented by Velleman for Kantian practical philosophy. Suitably revising the account for our purposes, we find that the parental relation explains 1. However, due to the paternalism problem, we find that it cannot explain 2. In the parental relation alone, we cannot come to regard ourselves as persons because we stand always under the laws of the parents—whereas persons are subject only to laws they give themselves (alone, or at least along with others).[7] So we attempt in the third section to solve the paternalism problem through engagement with the Piagetian account. We find that in acts of *co-legislation*, with peers, each party is recognized by the other as self-legislative, and in seeing this in the eyes of the other, each can recognize it of themselves; the peer relation will thus explain 2. But in denying positive contributions from parental authorities to moral development, Piaget leaves the basis of this co-legislative activity unexplained. We remedy this in reconciling the contributions of parents and peers—adumbrating the essay's original, positive picture of development.

Subsequent psychological literature has seen itself loosely divided along these lines—those who look to the parents and those who look to the peers.[8] The following offers a synthesis of the basic elements of the Freudian and Piagetian theories for a Kantian moral development. Arriving at developed moral personhood is necessarily a result of having had authorities who legislated to us and equals with whom we acted as co-legislators.

1. Developed Moral Personhood

An account of moral development attempts to outline the transition from infancy (an undeveloped state) to moral personhood (a developed one). To offer a complete account of development, we will have to be clear about our conception of moral personhood. Since this essay attempts to provide an account for Kantian practical philosophy, it helps itself to a certain Kantian

6 "Awakening to personhood" is Velleman's language, concerning our shared conception of the end state of development.

7 Cf. Kant, *Metaphysics of Morals*, 6:223 (hereafter parenthetically). All citations to Kant are to the Akademie numbers listed in the margins of most editions; all English translations are the Cambridge editions unless otherwise stated.

8 See note 5 above.

conception.[9] As exposited in the introduction to the *Metaphysics of Morals*, for Kant,

> a *person* is a subject whose actions can be *imputed* to him. Moral personality is therefore nothing other than the freedom of a rational being under moral laws (whereas psychological personality is merely the capacity for being conscious of one's identity in different states of one's existence). From this it follows that a person is subject to no other laws than those he gives to himself (either alone, or at least along with others). (6:223, emphasis original)[10]

Imputability is much like responsibility or accountability. To ask whether an act can be imputed to an agent is to ask about a sense in which she can be regarded as the source of the action; imputation is "the *judgment* by which someone is regarded as the author (*causa libera*) of an action" (6:227). With authorship in mind, we are led to conceptions of autonomy and personhood defined with respect to *self-legislation*: acting in accordance with laws (principles of action) we give ourselves. Self-legislation requires that we be able to conceive of two aspects or parts of ourselves, for Kant an empirical law-receiver and an

9 While Kant has a well-developed account of the person, one interpretation of which is exposited here, I do not take him to have a substantial account of the person's basic development. (Our narrative is not Kant's, though it is undertaken in a broadly practical-Kantian spirit.) Kant has much to say about how we may become *good* persons—more sensitive of and firmer in our commitment to duty (see, e.g., 6:477–84; his account of the cultivation of character in *Anthropology from a Pragmatic Point of View*, 7:292–95; and perhaps most topically, *Lectures on Pedagogy*)—but little to say about how we become persons *simpliciter*— that is, how we become morally culpable in the first place. Kant does think that before a person can be held responsible, they must understand certain "subjective conditions of freedom." I concur with Kate Moran that if what it takes to become a person is really education, "it must be understood more broadly than the kind of education that Kant outlines, for example, in *Lectures on Pedagogy*. The normal experiences that a person has in his or her formative years will serve as a sufficient education" (*Community and Progress in Kant's Moral Philosophy*, 162–63). I think we are better-off not calling it a form of education at all—one is not taught how to be assessable in terms of moral requirements. In a normal development, given the experiences this essay identifies, one develops a moral psychology satisfying the "subjective conditions of freedom" and ascends to personhood without being formally taught any special subject matter.

10 Note that "moral personality" is not exhaustive of the person (there is also psychological personality)—a person is a subject whose *actions* can be imputed to him, and one cannot be a cause to actions in the sensible world without also being considered as a sensible being (*homo phaenomenon*). The term 'moral personhood' in this essay refers not just to Kant's idiosyncratic use of moral personality as free subjectivity but to the more general condition of being a person, the two aspects as one. Longuenesse offers an interpretation of psychological and moral personality along just these lines ("Kant's Multiple Concepts of Person," 170).

intelligible law-giver. In natural persons, this is the distinction between *homo phaenomenon* and *homo noumenon*. We view ourselves under two attributes:

> First as a *sensible being*, that is, as a human being (a member of one of the animal species), and secondly as an *intelligible being*.… The senses cannot attain this latter aspect of a human being; it can be cognized only in morally practical relations, where the incomprehensible property of freedom is revealed by the influence of reason on the inner lawgiving will. Now the human being as a *natural being* that has reason (*homo phaenomenon*) can be determined by his reason, as a *cause*, to actions in the sensible world.… But the same human being thought in terms of his [moral] personality, that is, as a being endowed with inner freedom (*homo noumenon*) is regarded as a being that can be put under obligation. (6:228)

While differing in content, in form one may be reminded of Aristotle's dualistic conception of the soul—distinguishing between the rational and the nonrational: while the human being has reason, "they evidently also have some other [part] that is by nature something apart from reason."[11] The person, despite their consisting in two parts, is still unified, and the nonrational part "shares in reason in a way," that it "will have reason by listening to reason as to a father."[12] Similarly, *homo phaenomenon* is defined as a "natural being that has reason" in the sense of being "determined by reason." A Kantian cast of the two-part distinction is that of an ideal, intelligible *legislator* and that of a real, empirical *obedient*—their unity consisting in that which may be *self-legislative*.[13]

We see many of the famous tensions of Kantian practical philosophy rooted in this basic conception of what it means to be a person: mere animality and impulse on the one hand, rationality and lawfulness on the other.[14] Of course, infants and children, as we well know, are already impulsive—they are born a member of the animal species human being. One question for a theory of development thereby becomes: How is it that the infant eventually becomes a

11 Aristotle, *Nicomachean Ethics*, 20 (1102b).

12 Aristotle, *Nicomachean Ethics*, 20 (1103a).

13 Kant elsewhere defines the person less technically: "the body … constitutes, in conjunction with the self, a person" (*Lectures on Ethics*, 27:387). Here again, the person is a unity of itself as a sensible being (body) and intelligible being (self). We are concerned with the corresponding distinction in one's psychological life—the receptive, bodily part of the psyche and the legislative, intelligible part of the psyche.

14 The distinction is not in fact so simple. Our developed inclinations are "value-cognitive" due to the peculiar nature of our animality (the only animality that bears this form of relation to reason). Cf. Wood, "Feeling and Desire in the Human Animal." Nevertheless, I do not think this entails that our animality is *essentially* rational: it does not entail the total inability to distinguish our animality from rationality, as some, including Wood, have thought.

member of the rational, intelligible community of ends? If we assume that it is possible for such membership to never be attained, what types of experiences occasion this development?[15] What explains the relevant partitioning of moral psychological life into these two conceptually distinguishable parts, acquiring a "legislative" figure, a moral authority subjecting impulsiveness to lawfulness? How do we end up such that "morally practical reason in us *is* humanity (*homo noumenon*) that gives laws to us"?[16] It is this aspect of the person that just is constituted by the moral law: the true ideal of pure practical reason personified.

In coming to these questions as partially guiding questions of moral development, it is shown that Kantian theory already contains an understanding of the ideal of morally practical reason as a figure resembling that of the Freudian ego-ideal—the ideal of human personality as such. Though we will take the parental relation to answer the above questions, explaining the partitioning of our moral psychology into these two parts, it will not explain how we become a recognized *unity* of them, ultimately self-legislative, and thus in relations of reciprocal accountability.

The person, in their unity, is a member of a *real*, rational community, in its unity.[17] In such a community, one is called to justify one's actions before others, and one is called to hold the acts of others liable to assessment in light of principles. To qualify for membership, one must recognize oneself and be recognized by others as both legislator and obedient in the self-same entity, as self-legislative. Moral communities require conceiving of members as entities capable of both *giving* laws—a capacity necessary for participating in the cooperative process (for engaging in joint law making), and as entities capable of *receiving* and *realizing* laws—a capacity necessary for acting in accordance

15 Studies of so-called "feral" children support the assumption that such membership may never be attained—many children not exposed to social relations during the critical development period fail to ever acquire a language or to display moral reasoning/mature ego functioning. See especially Vyshedskiy et al., "Linguistically Deprived Children"; and Curtiss, *Genie*. See also note 22 below.

16 Kant, *Nachlass*, 23:338 (translation by Konstantin Pollok, via personal correspondence).

17 By the "rational, intelligible community of ends" previously, I mean something akin to Kant's ideal of a "kingdom of ends" acting under the idea of membership in a "systematic union of different rational beings under common laws" (*Groundwork for the Metaphysics of Morals*, 4:443). The kingdom of ends is controversial, though; I mean only to analogize. The idea of the "rational, intelligible community," like the kingdom of ends, is "merely possible"—it is an idea. The *real* moral community addressed here (a unity of the two other descriptions, that of the animal species and that of the intelligible community), by contrast, is not merely possible but really possible and indeed actual—it is constituted in part by an *ideal* of such a community, of rational beings relating in union, and it is in light of this ideal that we hold the nature of our relations to assessment. The end of development is membership in this *real* community.

with the principles jointly constructed by such a community. Finally, it requires identifying these members as *simultaeneously* the two, as those who not only receive the laws but who receive the laws *as their own*; this is what separates a free community of equals, a community of ends, from a despotism. This final requirement substantially expresses the necessity of recognizing the person as a unity. Here, I will argue, we need relations of equality; here we need the peers.

According to Kant, a person "acts under the *idea* of freedom."[18] On our more specific conception, to act under a recognition of oneself as a self-legislator (as a unity of law giver and receiver) is to "act under the *idea* of freedom."[19] Importantly, Kant claims, we are "really free in a practical sense" acting under the mere idea of freedom (we must act *as if* we were free), as we must hold the same practical laws to us as would hold for the actually free, and "thus we can escape here from the burden that weighs upon theory."[20] The conception of the person offered above is not centrally ontological but *practical*. In practical life, you do not treat yourself or others as merely bodily but also as abstractions, social identities, worlds of thought, and authors of action, among myriad other dizzying, unexplained sorts. It is also here that I avoid the burden that weighs upon theory: a genealogy of moral development does not have to demonstrate how we mystically become willers independent of causal chains but instead must demonstrate how we become beings who must *regard* themselves as autonomous willers—this recognition is enough for truly practical freedom.[21] It is

18 Kant, *Groundwork for the Metaphysics of Morals*, 4:448.

19 Rather, it is implicitly (and more inclusively) to act on conditions that suppose this recognition.

20 Kant, *Groundwork for the Metaphysics of Morals*, 4:448n.

21 This strategy hopefully suffices to avoid certain inscrutable metaphysical questions related to offering an account of Kantian moral development. For example, the grounds of freedom are inscrutable to theoretical cognition and are atemporal (Kant, *Religion Within the Boundaries of Mere Reason*, 6:138), yet, I want to say, offering an account of the development of a given individual to the age of autonomy is not helplessly obscure to us. This is because: (1) the end state of development is the state of "being really free in a practical sense," which epistemically requires only certain kinds of recognition; and (2) the theory of development is of a practical-psychological, speculative sort (relying on reflection, self-reports, and surveys, what is *communicable* about our practical experience) rather than a truly empirical-psychological sort, where we could expect to do something akin to mapping the development of freedom through fMRI scans. As Kant tells us, "Freedom—a property which is made manifest to the human being through the determination of his power of choice by the unconditional moral law—is no mystery, since cognition of it can be *communicated* to everyone; the ground of this property, which is inscrutable to us, is however a mystery, since it is not given to us in cognition" (Kant, *Religion Within the Boundaries of Mere Reason,* 6:138). Even if one thinks the grounds of freedom and morality are outside of time and explanation, the first- and second-personal practical recognition of oneself and of others as free, legislative beings, is not. I thank an anonymous reviewer for pressing this point.

clear that an infant cannot be identified with this condition. However, somehow along the way, at least some human beings develop into the kind of creatures who (at least implicitly) regard themselves as choosing courses of action on the basis of their own reasons and can be held to justificatory standards for those reasons. Somehow, we may develop into the kind of creatures that can have our actions imputed to us and that stand under (public) moral principles. Though we surely have some natural tendency toward this development, a theory of development will clarify some material conditions from which personhood arises *on the occasion*—i.e., it will clarify what must happen in experience before we are awakened to our personhood and the personhood of others.[22]

In our theory, we consider two sorts of relationships that developing children have—relations with authorities and relations with equals, whose contributions are understood as distinct. Further, in our narrative, authorities and equals take on special and evocative characters—the parents and the peers. The figure of the parent and the figure of the peer in this essay are reductive, they are not the same as a *given* parent or a *given* peer.[23] In the progress of normal development, an actual parent eventually becomes a peer. Nevertheless, especially earlier in development, an essential characteristic of a parent-child relation is that it takes the form of a relation of authority—"the parent" has a right to rule.

2. The Role of the Parents

The child, it is thought, is *raised* by the parents, and this has a transparently developmental connotation. In an important sense, parents have final say concerning what ends are set for the child and by what rules they can pursue

22 In the case of language acquisition, Chomsky contrasts "on the occasion of experience" with "derived from experience" (*Cartesian Linguistics*, 66–69, citing Cudworth, *Treatise Concerning Eternal and Immutable Morality*). The language faculty is innate on Chomsky's view but not naïvely innate, where somehow the faculty does not require exposure to certain empirical content for development. Rather, it "grows," "like a limb"; it is "something that happens to a child *placed in a certain environment*" (Chomsky, *Language and Thought*, 29). Many "feral children" never develop into language users—they lack the necessary experiences (exposure to language activity in the critical development period). It is similarly possible that a child misses necessary experiences and never becomes a fully developed moral person. Cf. Vyshedskiy et al., "Linguistically Deprived Children"; and Curtiss, *Genie*.

23 Not all children have actual parents. E.g., a child may be raised in foster care with no stable parental figure. Conversely, a child could be raised in a village with no other children. Still, someone or set of ones must play the functional roles of parent and peer, idealized authority and equal, for development to be possible. (Note minimal resources to diagnose *deficient* developments, where such figures are unstably present.)

them—two elemental features of practical life. It is unsurprising, then, that the Freudian tradition would think to center this relation.

2.1. A Rational Superego

On an orthodox Freudian view, psychoanalysis is concerned with means and not ends—it may explain where our motivations come from and equip us to do something about those perceived destructive motives we wish to reform or uproot, but it tells us nothing about whether those motives are moral or immoral, rational or irrational, in the Kantian sense. All motives of the organism's action originate in external stimuli and biological instinct (stimulus that arises from within the organism itself), where the two fundamental forms of instinct are self-preservative and sexual (in terms of which more "highly specialized" motives can be unproblematically dissected).[24] Here, considerations of autonomy and heteronomy or rationality and irrationality in the Kantian sense seem inoperative.

On an orthodox Freudian view, the critical account of "moral" development consists in our introjection of the authority of the parent. Introjection is a process by which one unconsciously internalizes representational images of others, and in the case of moral development, the child internalizes a picture of the parent (for Freud, especially the father) as a disciplining authority, an object of fear. The ego is impelled to introject the parent by the arational, sexual impulses of the id, who takes the parent as its first object.[25] Though the parent becomes desexualized when introjected, "their strength, their severity, their inclination to supervise and punish" remain.[26] The picture so internalized now lives in the child's head, as an independent agency and fixture of its psyche, the *superego*. When the child considers acting against the laws of this authority, so the toy story goes, they hear the voice of the father as the voice of conscience, commanding them against it; the ego and id's impetus to so act comes into conflict with this voice, and this produces feelings of anxiety and guilt. On an orthodox Freudian view, the formation of the superego has a destructive effect on the ego (and id)—the ego is burdened with conscience anxiety and consciousness of guilt, and it is now tasked to negotiate the "claims of the three agencies which it serves," itself, the id, and the superego—it has thereby introduced a new set of conflicts for itself to adjudicate.[27]

24 Freud, "Instincts and Their Vicissitudes," 118, 124; and Deigh, *The Sources of Moral Agency*, 129–30.

25 "[The] categorical imperative is thus the direct heir of the Oedipus complex" (Freud, "The Economic Problem of Masochism," 167).

26 Freud, "The Economic Problem of Masochism," 167.

27 Freud, "The Economic Problem of Masochism," 167.

Anxiety and guilt arise from the fact that the voice is not just of a disciplinarian to be feared; it belongs to someone the child wishes desperately to emulate. The child seeks to do as the father does. That is, the parent is not only internalized into the superego as disciplinarian, as an object of fear, the image of the parent is also internalized as the *ego-ideal*, an object of admiration.[28] But the child cannot be like the parent, the model they have is of an authoritative adult, and they are no authority and no adult; there is a necessary distance between the ego and its demanding ideal, the superego. The child lives in fear of an unsatisfiable ideal, an ideal they do not realize is of their own creation, one that stands to discipline them for failing to satisfy it; moral authority gets its first grip on us by way of the ego's apparently arational fear of its own (unconscious) creation.

The tone of this orthodox Freudian account intimates a moral life that is dark and forbidding and not exactly moral, in the Kantian sense. Among other revisionary work, one attempt to make use of the Freudian account for Kantian purposes is of special importance to this essay: Velleman's synthesis in a "A Rational Superego."[29] The success of such a synthesis, according to Velleman, depends on two conditions: "on Freud's side, that the ideals incorporated into the superego include an ideal of practical reason, and on Kant's side, that the categorical imperative—which is an ideal of practical reason—take the form of an ego ideal" (558). The latter condition having been met in the first section, Velleman's innovation consists in meeting the former—that is, in ensuring that the orthodox Freudian account may be "purged of its antirationalism" (558). Specifically, on the orthodox account, the "ideals" internalized into the superego are representations of actually existing social norms and conventions, "of social respectability or conventional propriety" (532). Velleman hopes to demonstrate that what is internalized can "contain an ideal of practical reason, an ideal of personhood as rational nature" (532).

On Velleman's view, when a child internalizes the authority of the parent, this contains an ideal of humanity—when the parent loves and sacrifices for the child, when they have concern for the child's needs aside from their personal interest, they thereby treat the child as an end in itself. It is this capacity that the child responds to in idealizing the parent; an essential part of what the child takes in is "nothing other than their practical reason . . . by which their immediate self-gratification is subordinated to rational requirements embodied in another person" (556). The ideals internalized thereby contain the categorical imperative, "which just is a description of the capacity to take persons as ends" (557).

28 Cf. Velleman, "A Rational Superego," 539–43; and Freud, "The Ego and the Id," 60.

29 Velleman's article will hereafter be cited parenthetically. Note that the exegetical dispute is beside our point—this essay is indifferent to whether the Kantian-compatible account is an alteration of Freud's views or internal to them. Cf. Longuenesse, *I, Me, Mine*.

We may internalize a figure of our parents as an ideal because as children we idealize our parents. In the earliest stages, though they are in fact imperfect authorities, we take their word as Word:

> Although the child may overvalue his parents as the noblest and loveliest specimens of humanity, he does not err in loving them, to begin with, as specimens of humanity, in the Kantian sense of the word. And when he later internalizes their tin nobility and paper loveliness, he must also internalize their humanity, which is pure gold—a standard not to be superseded by other ideals. (557)[30]

Velleman thinks the Freudian theory offers a valuable psychological model for three central features of Kantian practical philosophy. The first is the moral law's "dual status as a prescription and an ideal" (531). The moral law tells us what to do. It legislates, but it also describes what an ideal will does; its prescriptive force is justified by its status as an ideal. In presenting the rational superego as a special unity of the superego and ego-ideal, the rational superego tells us what to do, and it gives us a model to emulate (and the superego has authority by way of its status as an ideal). The second feature involves the interplay between the moral law's necessity and the fact that we give ourselves the law. How are we both bound by the authority of the law and ourselves the authority? How is one both legislator of and obedient to the selfsame law? The rational superego makes this "concretely imaginable." The authority of morality is represented as an internalized external authority, and "our ability to exercise moral authority over ourselves is explained by the familiar psychological process of internalizing other people" (531). The third feature is the "only plausible answer" to the question of moral development:

> Kantian ethics is an ethics of respecting persons, others as well as ourselves. *But what awakens us to the personhood of others, to the fact that the creatures around us are persons like ourselves?* Freud gives the only plausible answer to this question. The main theme of Freud's moral theory is that we are inducted into morality by our childhood experience of loving and being loved—the experience without which we would neither idealize nor internalize a parental figure. Love is our introduction to the fact that we are not alone in the world; and morality as formulated by Kant is our practical response to that fact. (532, emphasis added)

30 "In the beginning was the Word [*logos*], and the Word was with God, and the Word was God" (John 1:1). A consequence of idealizing the parent is taking their pronouncements as having special, approaching scriptural authority. "Well, my father said . . ." is the beginning of what seems to the child a genuine justification.

Velleman's theory of the rational superego provides a brilliant model of the first two features (the second with some qualification).[31] However, I think it falls short of explaining the third. The parental relation cannot alone explain how we are awakened to our own personhood and the personhood of others. Velleman's explanation is that the child is awakened to their own personhood in witnessing the parent treat them as an object of care, and the child is awakened to the personhood of others in seeing the parents as entitled to make authoritative claims on them. But recall that personhood does not rest in each of these separately—it rests in their recognized unity in the selfsame being (the relation of a personified moral authority and a morally considerable being that is not legislative but listens to the legislator "as to a father").[32] Here arises the paternalism problem.

2.2. *The Paternalism Problem*

If we are inducted into morality through the love of our parents, and our own moral authority is nothing other than an internalization of our parents' love and idealized authority, we can offer a theory of development that is compatible only with a moral paternalism, a "morality of heteronomy."[33] The parent does not relate to the child in the way that persons relate to each other; in relations of persons, each has a duty to respect the autonomy of the other, to respect the fact that the other is subject only to laws they give themselves (alone, or at least along with others), and so a duty to allow the other to legislate for themselves. Moral respect is ultimately a respect for the other as a moral equal, and the parent and the child do not stand in terms of equality. The parent's love for the child is love for a party more adequately described as a patient, one to whom the parent is called to give laws, and not for a person, one whose developed self-legislative capacity they would be called to categorically respect.[34] And conversely, the child loves and idealizes the parent as an authority, as a law

31 We may be seen as diverging from Velleman on the details of the "interplay of external and internal authority," as the essay will hold that the identification of moral authority with the image of the parent must fade, to be recognized as our own power.

32 Aristotle, *Nicomachean Ethics*, 20 (1102b).

33 Piaget's term of art in *The Moral Judgment of the Child*.

34 Indeed, considering Velleman's conception of love in "Love as a Moral Emotion" (love of an ideal of the other's person as rational nature), in conjunction with the functional role of love in his account of moral development, suggests a vicious circle: children acquire an ideal of the person as rational nature (in the parent's image) as a *response to* the parent's love, but the parent's love can be nothing other than love of an ideal of the person that the child must have *already had*. The parent's love cannot be both cause and effect of the child's being constituted by such an ideal.

giver, not as a whole person.[35] A person is not only one whose autonomy I am required to respect but also one who is required to respect my autonomy: a moral equal. The figure of the parent, insofar as it always must be someone capable of overruling me, awakens me not to the personhood of others and myself but to moral authorities (the parent) and moral patients (myself). The paternalism problem may be seen both in an analysis of this abstract external relation and in an analysis of the internalized psychical materials. We begin with the former.

If moral relations are relations of moral equality, it would be surprising that the child awakens to their moral personhood in a relation of inequality. A moral person, as we have said, acts under the idea of freedom, under the idea of themselves as both legislator of and obedient to that which governs them; the child, in the parental relation, is precisely not legislator of that which governs them—legislative power belongs to the parent. The rules of the parent do not require the say or participation or contribution of the child; the child may kick and scream, and the rules may remain justifiable. The child is often made explicitly aware of the fact that mutual justification is not needed ("because I said so …"). In such a relation, the child regards the authority over them as outside them and know themselves as an obedient.

Moral relations, as relations of equality, are founded more concretely on reciprocal respect and concern. The relation between the parent and the child does not seem to be one in which the child is capable of reciprocating what is offered by the parent, it is (as acknowledged by Velleman) just this *sacrificial* concern that captures the child's eye. When a parent cares for the infant at considerable self-sacrifice, the infant makes no offer; they do not reciprocally appeal to the interest of the parent. What the parent gains in the dedication of selfless service to the child is surely something immeasurable, but it is cheapened by the term 'interest', and in such service, what the parent stands to gain is precisely the thing beside the point.[36] The parent should offer their loving

35 In Freudian theory, what is internalized is not the parent in whole but the parent in their capacity as authority; the child internalizes *the parent's superego* (Freud, *New Introductory Lectures on Psychoanalysis*, 136). This is not exactly an ideal of personhood (as Velleman claims) but an ideal of *moral authority personified*. What the parent has, marking them as authorities rather than whole persons, is a recognized *right to overrule*. If, in certain moments, the tempestuous toddler refuses to recognize this right, it is usually born of impulse (the idealization of the parent also rationalizes their authority), and *that* there is a recognition of the right is certainly clear to the child, who sees the whole social order enforcing this right.

36 The intention of the given parent may not be so selfless—what matters is what the child is capable of seeing and idealizing. Even if the nobility is tin and the loveliness paper, to the child it is pure gold.

care expecting exactly nothing from the infant (as infant) in return. Of course, as development proceeds, children increasingly owe their parents—gratitude, the completion of chores, displays of love, and so on. This is, however, twinned with the gradual effacing of what is essential to the parent-child relation as an authority relation.[37]

Velleman argues that the parent's sacrificial concern for the child treats the child as an end in itself, and the child thereby internalizes the categorical imperative, fully formed, in their internalization of the parental figure. But concern aside personal interest is not enough to show full treatment of *persons* as ends in themselves. I cannot impose my will over that of another adult because I think it in their interest, even if I am sacrificing my own interests. To treat persons as ends in themselves, I must respect their autonomy, their capacity to set ends for themselves. Kant calls the capacity of humanity to "set ... itself an end" an end in itself.[38] By contrast, the parent often has a duty to treat the child paternalistically—not to respect their capacity to set ends for themselves (because this capacity is not yet developed) but to guide development such that the child may eventually govern themself by moral principles. Because the parent loves the child as a patient, a developing being, and not as a developed person, the parent's love cannot be the vehicle of respect needed for the child to awaken to their own personhood.

One may think: Is this future-respecting concern not itself respect for the child's capacity to set ends for themself? While I think this confuses the modality of *capacity*, even if we take the premise of the question for granted, I think we must still answer negatively.[39] It may appear to be respect for the child's capacity for self-legislation from the parent's perspective, but what is needed is a description from the child's perspective, as they are the party whose experiences are necessary for the internalization of the categorical imperative. And here I think it is safe to say: paternalism appears to the child as just that—paternalism. They see themselves as overridden; they may kick and scream, but the rules remain nonetheless. In the course of development, the parent must give the child space to strike out on their own and make their own mistakes; increasingly also, they will be owed reasons for rules. But especially in the earliest stages, what matters is that in this relation, the parent always has *overridingness*; the child must put on the seat belt, must hold hands while crossing the street—about all of this, there is little to nothing to negotiate.

37 As we find in the next section, relations of authority, in normal development, give way to relations of equality, and eventually, the given parents themselves will become peers.

38 Kant, *Groundwork for the Metaphysics of Morals*, 4:437.

39 Here, capacities (which one may be able to exercise now, *ceteris paribus*) and developmental potentials should be held apart.

Now, we should not mistake the baby for the bathwater. The rational superego remains an important psychological model for the structure of internal law giving: the structure of the authority of *homo noumenon* over *homo phaenomenon* may even be described as a paternalistic structure (and is so described in the Aristotelian analogy) and so rendered explicable in its rudiments by the internalization of parental authority over the child. *Homo noumenon* begins as the figure of the parent, an ideal and a disciplinarian who gives laws, and *homo phaenomenon* the child, a receiver of laws, subject to rules and obliged to follow them. The child internalizes a picture of parental authority over them, in which an idealized disciplinarian gives laws, which a natural being stands under. The relation becomes, in the course of development, the procedure of internal law giving, in which we conceive of two aspects of ourselves that make self-legislation possible. The relation is not one of mere power—the child, in loving the parent, idealizes them, and power becomes *authority*. This is the embryo of practical normativity.

Even when, however, we look more specifically to these internalized psychical materials, we still have not quite arrived at a psychology characteristic of personhood. For the child, the internalized authority figure is not seen as an aspect of themself; the internalized authority is an image of the parent, an introjection bearing their mark, speaking in their voice. The child is equipped to subdue their own animality and so has acquired some of the relevant psychical materials for self-legislation—they are capable of being disciplined subjects to laws, but they regard this subduing as issuing not from their *own* authority but from without, from the parent. The child's laws are still in content and form the parent's laws, and in the structure of internal law giving at this stage, the legislator is still claimed by the external parental figure.[40] Sole obedience to the external is nothing other than heteronomy. Somehow, the child will have to come to recognize the authority as their own and thus not remain in obedience to the laws of others but genuinely impose laws on themselves. The materials provided by the parent will have to be correspondingly transformed. In one minimal sense, the child will have to kill the father—in their practical suppositions the child must conceive of themself as the legislator of their own principles of action; this is what it will mean to regard ourselves as autonomous willers, to awaken to our personhood.[41] For that, we will need a different sort of relation.

The parent's core contribution to this account of moral development is thus not in explaining our status as imputable but is rather in equipping us

40 Even the process of internalization and idealization appears passive and unconscious, a feature of the authority relation. Only when we begin to reach the age of moral maturity, capable of de-idealizing the parents, are we even positioned to see that the construction of the superego was in part the ego's doing.

41 See again Kant, *Metaphysics of Morals*, 6:223.

to subject the childish aspect of ourselves to authoritative rule (abstracting from the question of *to whom* rule is identified) and so how to be disciplined law receivers.[42] We have now unconsciously internalized the parental figure as the authoritative aspect of our moral psychology. What we have not done is recognize ourselves or others as a real unity of authority and obedient (claiming the former as our own), elevating us, as whole persons, to membership of a real, rational community, one in which we are really held liable to assessment by our fellows. When the child's laws are still, in the end, the parent's laws, from the second- and third-personal perspective, it is the parent, not the child, that is regarded as in-the-end culpable; and from the first-personal perspective of the child, they themself remain (legitimately) subject to the parent's laws. Paternalism, necessarily a treatment regarding the child as less than fully self-legislative, prevents the child from recognizing its own rational nature, its rational nature as its own, and so the mirror must be held to the child in another way.

3. THE ROLE OF THE PEERS

When a child interacts with peers (here represented by other children) outside of the purview of authorities, the child engages with others by their own power, on their own terms. Relations between equals, between peers, can embody values of cooperation, reciprocity, and mutual respect. Piaget considers this insight scientifically and systematically in *The Moral Judgment of the Child*.[43]

42 Recall that the empirical literature finds "inductive discipline, authoritative parenting, responsiveness, and involvement" associated with "moral maturity" and "mature ego functioning." See Walker et al., "Parent and Peer Contexts for Children's Moral Reasoning Development"; and note 5 above. Proper ego functioning is predictive of moral development over time in a particular way: "mature ego functioning on the part of socializers scaffolds children's moral development by engendering an affective context in which ... emotions are *expressed appropriately*, and *self-control* is evident even in difficult circumstances" (1036). Cf. Hart et al., "Overt and Relational Aggression in Russian Nursery-School-Age Children"; and Matsuba and Walker, "Moral Reasoning in the Context of Ego Functioning." The parent's functional impact, in our account, is a kind "scaffolding" that renders possible other forms of social associations—those with peers.

43 Hereafter, *The Moral Judgment of the Child* is cited parenthetically. Note that Elliot Turiel has provided a critical appraisal of Piaget (and Kohlberg) in light of contemporary evidence: "A large body of research has, on the one hand, provided very strong evidence that the type of differentiations proposed by Piaget and Kohlberg do not adequately characterize the development of moral judgements, and, on the other hand, *strongly supported* fundamental propositions put forth by Piaget and Kohlberg regarding the construction of moral judgements of welfare, justice, and rights through reciprocal interactions with multifaceted aspects of the social environment" ("Thought, Emotions, and Social Interactional Processes in Moral Development," 24). In the context of Turiel's paper, it is clear that the evidence against Piaget is evidence against specific empirical claims of stage development

3.1. *The Moral Judgment of the Child*

Piaget was a Kantian, and so we fortunately do not have to do the work of reconciling his premises with Kantian ones. While orthodox Freudian theory considers children as passive internalizers of already existing social norms, Piaget considers children as active participants in the construction of norms. Genuine moral norms are properties that arise from certain forms of social interaction, those of equality and cooperation. Piaget distinguishes relations of "constraint" and "cooperation":

> The great difference between constraint and cooperation or between unilateral respect and mutual respect, is that the first imposes beliefs or rules that are ready made and to be accepted *en bloc*, while the second only suggests a method—a method of verification and reciprocal control in the intellectual field, of justification and discussion in the domain of morals. It matters little whether this method be applied immediately to all the rules imposed by the environment or only to one aspect of behaviour: once it has come into existence it has the right to be applied to everything. (91)[44]

Recasting the paternalism problem in Piagetian terms: the relation of the parent and the child is a relation of constraint and unilateral respect, and thus it never establishes the method of justification in the domain of conduct. For that, we require cooperative relations—relations of peers.

Piaget's discussion of the peers involves an insightful study of games. Children's participation in games opens up for them a field of norms over which they have control—norms applicable to themselves and other children with whom they play, applicable on conditions of mutual respect; the rules of the game are not rules of the game if participants do not recognize them. Games thereby model at least three critical aspects of moral life: action in accordance with norms; the construction of norms; and the justification of norms that one considers valid. Children must play in accordance with the rules of the game, where they are held liable to assessment by others in light of these rules. The rules of the game are in part a result of deliberation—often children do not

(e.g., the egocentrism of Piaget's first stage is at tension with observations of early prosocial attitudes); it is not an indictment of the general spirit of Piaget's project. As indicated, the claim that judgments of justice and right arise through reciprocal interaction has only been substantiated.

44 Piaget continues: "Mutual respect is ... the state of equilibrium towards which unilateral respect is tending when differences between child and adult ... are becoming effaced; just as cooperation is the form of equilibrium to which constraint is tending in the same circumstances" (91).

know the rules entirely and so must invent and reform them together, or they know them and take up the project of altering them. When children engage in these deliberative procedures, they justify their rule proposals to others and challenge justifications that seem against their interests or not quite right.[45] The playground is the child's primitive moral community, a superficially supervised space of reasons. Studying the playground, Piaget draws a distinction between "constitutive" and "constituted" rules:

> There are two kinds of rules, those that are constitutive and render possible the exercise of cooperation, and those that are constituted and are the result of this very exercise. . . . The rules of the Square, of the Coche, etc. . . . are "constituted" rules, due to mutual consent, and capable of being altered by general opinion. The precedence given to justice as opposed to chance, on the other hand, of effort over easy gain are "constitutive" rules, for without this "spirit of the game" no cooperation would be possible. (92–93)

The notion of *constitutive rules* here is distinct from its contemporary epistemological usage. In the contemporary sense, a constitutive rule of a game is just the rule that makes the game that game (that is, what Piaget calls "constituted rules")—if we move the rook diagonally in a game of chess, we are not playing chess. Nonetheless, these rules get an important kind of grip on us—there is force to the injunction "We're playing chess, so don't move the rook diagonally." The need for stability of such constituted rules is no match for the awesome imaginations of children at play. Children may be seen "playing chess" wherein the king moves two spaces instead of one. Nonetheless, there remain rules that get a kind of inalterable grip on the children's psyche, the second order sort, those that constitute and govern the deliberative, cooperative activity by which such alterations and formations are made—the rules without which we are no longer even playing *together*.

Developed moral rules in social life seem to take an analogous form to those of children's games: rules concluded through deliberation and that require mutual consent, as well as rules that are necessary conditions of this

45 Recall that the empirical literature suggests that discussions of dilemmas between children are more transactive than with parents (Walker et al., "Parents and Peer Contexts for Children's Moral Reasoning Development," 1034, citing Kruger and Tomasello, "Transactive Discussions with Peers and Adults"; and Kruger, "The Effect of Peer and Adult-Child Transductive Discussions on Moral Reasoning"). Parental discussions are highly transactive when they are "highly representational interactions ['socratic style . . . drawing reasoning out through appropriate probes, i.e. in cooperative, recognitive forms']," when parents play a more peer-like pedagogical role. By contrast, "highly informative interactions [opinionated lecture style] were associated with slow rates of moral growth" (1045).

deliberative activity. If my spouse and I speak at some length to establish the extent of privacy, and each agree never to read the other's journal, we have formed a constituted moral rule, one that would be morally dubious to break. This rule could be otherwise, but it was generated in a deliberative space constituted by rules that could not be—those constituting mutual, reciprocal respect—constitutive moral rules. Less solid in genealogy, cultural mores may be constituted moral norms. What precisely it means to fail to respect the dignity of another may be vastly different in content in Colombia than in Japan. It may be offensive in some circumstance, even morally, for an Italian to brush the underneath of their chin with an outward sweep of the fingertips of one hand, and yet be a completely meaningless gesture elsewhere. Nonetheless, the constitutive moral rule, to respect the dignity of others, remains constant. The "method of verification and reciprocal control in the intellectual field, of justification and discussion in the domain of morals," constitutive of relations of cooperation, will come to have "the right to be applied to everything" (91). The moral law, for Piaget, is the ur-constitutive rule of cooperative activity—the method of justification in the domain of conduct.

Piaget then confronts a problem: "how can these constitutive rules be regarded as themselves the outcome of mutual respect since they are necessary for the latter's formation?" (93). His answer is to wave it away: "The difficulty here is purely formal. Between mutual respect and the rules which condition it there exists a circular relation analogous to that which holds between organ and function" (93). It soon becomes clear why Piaget must dissolve rather than solve the problem: "Since cooperation is a method, it is hard to see how it could come into being except by its own exercise. No amount of constraint could determine its emergence. If mutual respect does derive from unilateral respect, it does so by opposition" (93). Piaget does not have much to identify by way of necessary conditions for the emergence of cooperation because here he refuses positive, necessary contributions from relations of unilateral respect (i.e., from the parents).[46] Instead, he contends, through a kind of spontaneous and self-reflexive process, cooperation both emerges and creates the conditions for its emergence—the "equilibrium-based solution."[47]

Now, I struggle to see how the "circular relation" in question could be anything but a vicious one: if only cooperation could create the conditions for its own emergence, then such conditions would never be created, for want of cooperation. The appeal to circularity appears to be a way of pulling rabbits

46 His position on the parents in general (rather than specifically their relation to the development of mutual respect) is more developed than I spend time with here. See Piaget, *The Moral Judgment of the Child*, ch. 2.

47 Piaget, "The Relationship Between Morality and Law," 161.

from hats, and indeed, I think it leads Piaget to try to reach for a rabbit himself: the ability of cooperation to spontaneously and self-reflexively develop leads us to the conclusion that "two of three individuals who have lived their entire lives on a desert island will necessarily come up with the idea of justice, without implying that they had it in them to start with."[48] The idea of justice, what Piaget calls the "most rational of all moral norms," will spontaneously emerge within social relations, even between those who have not been taught or loved or raised, and thus all that is *really* necessary for moral development is the peers. This seems almost obviously wrong. It is not just the individual who had to develop to the age of morality, it is also the society.[49] We learn from John Dewey that the path to moral social life was surely long and torturous, and required the communication and transmission of "ideals, hopes, expectations, standards, opinions, from those members of society who are passing out of the group life to those who are coming into it."[50] Rational interaction was not always a constitutive ideal of social interaction. The young do not have the capacity to survive on a desert island even physically without caring authorities—"how much more then, is this the case with respect to all the technological, artistic, scientific, and moral achievements of humanity!"[51]

Piaget comes to this rather perplexing position on the relation of constraint to cooperation by first posing a very difficult problem for any developmental account that starts from our premises. Indeed, Piaget considers a conception of the parental contribution that is, in summary, quite similar to our own, that of the first stage of J. M. Baldwin's theory.[52] Here too, the child acquires conscience in an idealized parental image, becoming accustomed "to the presence of something in him that represents his father, mother, or in general the *law-giving personality*."[53] The trouble is, as Piaget puts it, having thus begun, "one cannot escape the following contradiction—namely, that rational logic is itself derived from social processes from which it is supposed to free itself in matters of moral psychology" (398). This is the central challenge set by Piaget:

48 Piaget, "The Relationship Between Morality and Law," 161.

49 Piaget's implicit premise is that a nascent society of two would not require such development.

50 Dewey, *Democracy and Education*, 3.

51 Dewey, *Democracy and Education*, 3. Consider again "feral children." For language acquisition, children require early exposure to language activity. Two infants on a desert island, with no exposure to such activity in the critical development period, will not spontaneously generate a fully formed language.

52 See Baldwin, *Social and Ethical Interpretations in Mental Development*.

53 Baldwin, *Social and Ethical Interpretations in Mental Development*, 49 (quoted in Piaget, *The Moral Judgment of the Child*, 395–96).

in our story, "morality begin[s] with conformity and end[s] in autonomy. . . . To account for this evolution," we must have an explanation that "really suffices to make us understand why constraint gives way to cooperation" (398). Everyone else, Piaget thinks, has failed to do so, instead giving "priority to the relation of constraint over against the relations of cooperation" (398). Piaget bites the bullet on this problem: in the end, constraint does not give way to cooperation by providing its original necessary conditions; rather, cooperation is the immanent equilibrium of all social relations. (If there are ever noncooperative relations, their inherent instabilities will eventually lead back to cooperation.)

In my view, Piaget thus errs in just the opposite way, giving undue priority to relations of cooperation. For Piaget, the relation of constraint and its characteristic "morality of heteronomy" is something merely to be overcome; for this essay it is instead, to borrow a Hegelian expression, something to be *sublated*. Cooperation cannot nebulously produce its own necessary conditions but emerges *only* on the basis of necessary contributions to development provided by relations of authority. We *can* give an account of how constraint gives way to cooperation, and we *can* escape the contradiction that rational logic is itself derived from social processes from which it is supposed to free itself in matters of moral psychology. And in bringing together parents and peers, we will show now how this is concretely imaginable.

3.2. *Parents and Peers*

In the joint construction of norms, *co-legislation*, we are awakened to and recognize the personhood of others, a respect owed to them and their autonomy. We know that when we stand in terms of equality, we cannot treat another as an obedient of our laws; we cannot treat them as we have learned to treat our animality. Any childish effort to assert our own principles over them is in vain, as we soon learn that they can do otherwise without impediment. In acts of co-legislation, we are also awakened to and recognize our own personhood. We see, in the reflection of the gaze of our co-legislators—our peers—a respect for our own personhood. This is a respect that we could not have before seen in the eyes of authorities, who did not require our consent or participation for the construction of rules, who always in some part had the reins on governance and an intention for guidance. The relation of co-legislation is thus our first relation of moral equality. Co-legislation brings forth the recognition of self-legislation within and toward both parties.

The analogous problem to Piaget's conditions of cooperation is this: Do we not already have to be *in some sense* self-legislative to participate acts of co-legislation? We will not wave this problem away, but answer it: the contribution of the parents has provided what is necessary to participate in acts of co-legislation.

The imposition of rule and discipline, and the internalization of an idealized parental authority, has made discursive, rule-governed activity with the peers possible; the child has an efficient power, even if they do not know it as their own. We can participate in such activity prior to a recognition of ourselves as legislative because peers recognize the moral authority in us as our own before we do, and we simultaneously return this favor. Two similarly situated children can participate in acts of co-legislation prior to each recognizing the legislative power in them as their own, because they can each see the legislative power of the other as belonging to the other. Finally, the reciprocal recognition of co-legislative activity allows each child to see, in the eyes of the other, a regard for their own legislative power, and in so seeing, recognize it of themself.

In order to participate in, approach, or even consider acts of co-legislation, we must come to the table with some capacity or constitution. And this does not materialize from thin air; it is not within the infant fully formed; it is disciplined and trained by the idealized authority of the parents, and this authority is internalized. If children came to the schoolyard with no authority over the impulsive aspect of themselves, they could not participate in any such joint deliberation; their consent to the rules of the game would not matter because they would be incapable of giving it; they may be excluded from the beginning or at the point of older or more developed children's recognition of these facts.[54] It is the parent who has a duty to legislate to the young child, as Kant claims in the *Pedagogy*, the child would otherwise be trapped in its "inner wildness," its animal nature (*tierische Wildheit*).[55] It is only by the parent's discipline that the child learns to self-discipline.

But without the peers, the child is not quite yet an autonomous person. For the child to be autonomous, they need authority over their own powers, but in the parental relation, it is still the parent who claims this authority. This is captured already by Piaget: in the first stage of the development of the child's

54 We can now offer a narrative tracking some of the empirical literature's suggestions about the different functional impacts of parental and peer contributions: early inductive discipline and authoritative parenting tends to "mature ego functioning on the part of socializers" thereby "scaffolding children's moral development" allowing for social interactions in which "emotions are expressed appropriately and self-control is evident even in difficult circumstances"; these contributions enable social interactions with peers that are highly transactive for moral development, generating advances in "higher moral reasoning" (Walker et al., "Parents and Peer Contexts for Children's Moral Reasoning Development"). Eventually the most transactive modes of engagement for parents are those simulating equality, and the parental relation itself gradually effaces itself into something peer-like. See notes 5, 42, and 45 above.

55 Kant, *Lectures on Pedagogy*, 9:442, 9:444, 9:449.

sense of justice, the idea of justice is subordinated to adult authority—it is no different from obedience: if the adult says it, it is just.[56] But in the end,

> the sense of justice … requires nothing more for its development than the mutual respect and solidarity which holds among children themselves. It is often at the expense of the adult and not because of him that the notions of just and unjust find their way into the youthful mind.… In contrast to a given rule … such as the rule of not telling lies, the rule of justice is a sort of immanent condition of social relationships or a law governing their equilibrium. And as the solidarity between children grows, we shall find this notion of justice gradually emerging in almost complete autonomy. (198)

Though we diverge from Piaget on there requiring "nothing more for its development" than the peers, this reveals the way in which the child suitably transforms the materials provided by the parent into genuine autonomy—solidarity among the children themselves, often against and at the expense of the adult! Children keep secrets among themselves, they resolve to break rules and not tattle to the adult, they coordinate to play tricks on the authorities, etc. The development of autonomy often appears as a *rebellion* against external authority. But why be solidaristic, and why rebel?

Children originally bring their parents' principles to the schoolyard, but they are challenged by the principles of others (even those of similarly situated children, whose principles are also merely inherited). This challenging exposure to the differing principles of others reveals the fault in our assumption that our parents are the "noblest and loveliest specimens of humanity," in at least our assumption that their word is Word, and so we may depart from them not from mere impulse but from genuine reasons—principled differences of our own. As the ideality of the parent was that which, for the child, licensed their authority, the impetus to rebel is thereby more particularly brought on by an understanding of the parent's imperfection; when the nonideal nature of the parent is thoroughly understood, the child begins to take legislation into their own hands. Because the challenge comes from equals, parental authority is not merely supplanted by the external authority of another—the challenge is not an imposition but a disputation. This recognition of the nonideal nature of the parent thus also motivates solidarity between children: when children bring their parents' principles

56 The "sense of justice" is for Piaget "the most rational of all moral notions" and must be the result of cooperation. It is here that we "penetrate into the child's actual consciousness" rather than the merely "juridical"—that associated with "a morality of constraint or of heteronomy" defining relations of unilateral respect and authority (*The Moral Judgment of the Child*, 195).

against each other, the children are reciprocally destabilized, and they thereby find themselves in a common situation—subjected to authorities and to rules they now find imperfect, subjected regardless of their imperfection, and in turn in search of principles they can really call their own. This shared developing-dominee-like status encourages the solidarity in which children support each other in the effort to part ways from the parent, to stand on their own two feet (not in solitude, but in relation to equals). The good parent seeks for their child the same—the effacing of the parent-child relation *is* its realization.

The turn from constraint to cooperation, from authority to equality, and the escape of the apparent contradiction of morality "deriv[ing] from social processes from which it is supposed to free itself in matters of moral psychology" consists in something of this process, a process of idealization and de-idealization. The first relation transformed power into authority, and idealization conferred legitimacy on the authority's laws; the second relation weakened idealization of the external and turned it inward, where each is recognized as exercising authority of their own, and where legitimacy now depends on the authority of each. The unilateral authority relation was imbued with a truly moral-developmental character after the process of idealization, when power became authority. The turn from constraint to cooperation was therefore not an inexplicable transformation from inert power to genuine moral relations but a relocation of the ideal to its proper place—not a named, voiced conception of an external figure but an ideal conception recognized as in part constituting ourselves and others, equals with whom we act and live together—an ideal in light of which we are now licensed to reciprocally hold each other to account.[57]

We gain autonomous authority over ourselves when we transcend the yoke of parental authority, when we recognize the laws that govern us as our own, when we recognize ourselves as legislative.[58] But as we require the parents to

57 So we propose, as an alternative to Piaget's equilibrium emergence of cooperative relations: (1) seeing the child as idealizing and internalizing a legislative power in the parental image (one which, by this idealization, rules with legitimacy); (2) seeing the necessary conditions for the emergence of co-legislative activity as having this efficient power, even if the child themself has not yet recognized the power as their own (they are able to do so with a similarly situated child because each can see the other as having this legislative power, even if each does not yet know it of themself); and (3) seeing that the reciprocal recognition of co-legislative activity results in each child being able to see in the eyes of the other a regard for each other's legislative power as belonging to each, and so finally being able to see it of themself. The resources Piaget lacked that allow us to account for authority giving way to equality are, first, the idealization proper to the parental relation and, second, the logic of reciprocal recognition proper to the peer relation.

58 Freud partially anticipates this, as "the course of childhood development leads to an ever-increasing detachment from parents." But for Freud, the de-idealization story is still one of authorities: "To the *imagos* [of the parent] they leave behind there are then linked

acquire the power to subject impulsiveness to lawfulness, we have the capacity to rebel against external authority *only by first internalizing it*. We turn away from an idealized external authority through the embrace of equals, who do not impose their ends on us, but by the conditions of co-legislative activity, respect our capacity to set ends for ourselves, and therein we claim authority over our own powers. True autonomy is the result of both the internalization of parental authority and a rebellion against it.[59] Only with authorities do we learn to take authority over ourselves, and only with equals do we learn that authority is our own.

4. CONCLUSION

Thus, we find that, in part by way of synthesizing the Freudian and the Piagetian, the roles of the parents and peers are not only compatible but complementary, and both needed for moral development. With parents, we learn to be disciplined law receivers; with peers, we learn to be law givers; and in their combination, we become *persons*, recognizing ourselves as legislative and obedient of the selfsame laws. Thus acting under the idea of freedom and relating ourselves to determining grounds of an entirely different sort—earning our membership in the real, rational community of ends.[60]

Stanford University
grantray@stanford.edu

the influences of teachers and authorities, self-chosen models and publicly recognized heroes, whose figures need no longer be introjected by an ego which has become more resistant. The last figure in the series that began with the parents is the dark power of Destiny.... There is little to be said against the Dutch writer Multatuli when he replaces the μοῖρα [Destiny] of the Greeks by the divine pair Λόγος και Ἀνάγκη [Reason and Necessity]" ("The Economic Problem of Masochism," 168).

59 This insight foreshadows the teenage years. In the United States, we do not hold children fully responsible until the age of eighteen. This is also usually the culmination of the teenager's rebellion against authority, when they no longer follow their parent's every rule, but resolve instead to create and follow more exclusively their own rules. Rebellion against parental authority is usually also in coordination with the peers. They resolve to cut curfew together, they resolve to experiment with drugs together, they resolve to drink wine and read Goethe in the woods, and so on. Rebellion is a rather dramatic (and characteristically American) expression of the movement. All that is necessary is the fading away of the idealization of the parents—their authority and judgment has to be seen for what it is, less than perfect, never quite what it could be.

60 I am especially grateful to Elizabeth Anderson, as this essay was originally written as an undergraduate thesis under her supervision. Her approach to the student's autonomous development demonstrated that authority and equality might be properly mediated in *pedagogy*. I am also thankful to many other peers, or perhaps my betters pretending to

REFERENCES

Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. 3rd ed. Hackett, 2019.

Baldwin, J. M. *Social and Ethical Interpretation in Mental Development*. Macmillan, 1897.

Boyes, Michael C., and Sandra Allen. "Styles of Parent-Child Interaction and Moral Reasoning in Adolescence." *Merrill-Palmer Quarterly* 19, no. 4 (1993): 551–70.

Cassirer, Ernst. *Kant's Life and Thought*. Yale University Press, 1981.

Chomsky, Noam. *Cartesian Linguistics*. Harper and Row, 1966.

———. *Language and Thought*. Moyer Bell, 1993.

Cudworth, Ralph. *Treatise Concerning Eternal and Immutable Morality*. Andover, Gould, and Newman, 1838.

Curtiss, Susan. *Genie: A Psycholinguistic Study of a Modern Day "Wild Child."* Academic Press, 1977.

Deigh, John. *The Sources of Moral Agency: Essays in Moral Psychology and Freudian Theory*. Cambridge University Press, 1996.

Dewey, John. *Democracy and Education*. Free Press, 1916.

Dunton, Kathryn J. "Parental Practices Associated with Their Children's Moral Reasoning Development." PhD diss., Stanford University, 1989.

Freud, Sigmund. "The Economic Problem of Masochism." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 19 (1923–1925), translated and edited by James Strachey. Hogarth Press, 1961.

———. "The Ego and the Id." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 19 (1923–1925), translated and edited by James Strachey. Hogarth Press, 1961.

———. "Instincts and Their Vicissitudes." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 14 (1914–1916), translated and edited by James Strachey. Hogarth Press, 1957.

———. "New Introductory Lectures on Psychoanalysis." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 22 (1932–1936), translated and edited by James Strachey. Hogarth Press, 1964.

Hart, Craig H., David. A. Nelson, Clyde Robinson, and Susanne Olsen Roper. "Overt and Relational Aggression in Russian Nursery-School-Age Children: Parenting Style and Marital Linkages." *Developmental Psychology* 34, no. 4 (1998): 687–97.

play the part: Lanier Anderson, Sarah Buss, Zane Nayfeh, Alexander Pereira, Konstantin Pollok, Debra Satz, Leif Wenar, and Allen Wood.

Hart, D. "A Longitudinal Study of Adolescents' Socialization and Identification as Predictors of Adult Moral Judgment Development." *Merrill-Palmer Quarterly* 34, no. 3 (1988): 245–60.
Hoffman, M. L., and H. D. Saltzstein. "Parent Discipline and the Child's Moral Development." *Journal of Personality and Social Psychology* 5, no. 1 (1967): 45–57.
Kant, Immanuel. *Anthropology from a Pragmatic Point of View*. In *Anthropology, History, and Education*, edited by Günter Zöller and Robert B. Louden. Cambridge University Press, 2007.
———. *Groundwork for the Metaphysics of Morals*. In *Practical Philosophy*, translated and edited by Mary J. Gregor. Cambridge University Press, 1996.
———. *Lectures on Ethics*. Edited and translated by Peter Heath. Cambridge University Press, 1997.
———. *Lectures on Pedagogy*. In *Anthropology, History, and Education*, edited by Günter Zöller and Robert B. Louden. Cambridge University Press, 2007.
———. *Metaphysics of Morals*. In *Practical Philosophy*, translated and edited by M. J. Gregor. Cambridge University Press, 1996.
———. *Nachlass*. In *Gesammelte Schriften, Band XXIII*, edited by G. Martin. De Gruyter, 1955.
———. *Religion Within the Boundaries of Mere Reason*. Edited and translated by Allen Wood. Cambridge University Press, 2018.
Kruger, A. C. "The Effect of Peer and Adult-Child Transductive Discussions on Moral Reasoning." *Merrill-Palmer Quaterly* 38, no. 2 (1992): 191–211.
Kruger, A. C., and Michael Tomasello. "Transactive Discussions with Peers and Adults." *Developmental Psychology* 22, no. 5 (1986): 681–85.
Longuenesse, Beatrice. *I, Me, Mine: Back to Kant, and Back Again*. Oxford University Press, 2016.
———. "Kant's Multiple Concepts of Person." In *Kant on Persons and Agency*, edited by Eric Watkins. Cambridge University Press, 2017.
Matsuba, M. K., and L. J. Walker. "Moral Reasoning in the Context of Ego Functioning." *Merrill-Palmer Quarterly* 44, no. 4 (1998): 464–83.
Moran, Kate. *Community and Progress in Kant's Moral Philosophy*. Catholic University of America Press, 2012.
Piaget, Jean. *The Moral Judgment of the Child*. Translated by Marjorie Gabain. Free Press, 1948.
———. "The Relationship Between Morality and Law." In *Sociological Studies*, edited by Leslie Smith. Routledge, 1995.
Scheffler, Samuel. *Human Morality.* Oxford University Press, 1992.
Turiel, Elliot. "Thought, Emotions, and Social Interactional Processes in Moral Development." In *Handbook of Moral Development*, edited by Melanie Killen

and Judith G. Smetana. Lawrence Erlbaum Associates Publishers, 2006.

Velleman, J. David. "Love as a Moral Emotion." *Ethics* 109, no. 2 (1999): 338–74.

———. "A Rational Superego." *Philosophical Review* 108, no. 4 (1999): 529–58.

Vyshedskiy, Andrey, Shreyas Mahapatra, and Rita Dunn. "Linguistically Deprived Children: Meta-Analysis of Published Research Underlines the Importance of Early Syntactic Language Use for Normal Brain Development." *Research Ideas and Outcomes* 31 (2017): article e20696. https://doi.org/10.3897/rio.3.e20696.

Walker, Lawrence J., Karl Hennig, and Tobias Krettenauer. "Parent and Peer Contexts for Children's Moral Reasoning Development." *Child Development* 71, no. 4 (2000): 1033–48.

Wood, Allen. "Feeling and Desire in the Human Animal." In *Kant and the Faculty of Feeling*, edited by Kelly Sorensen and Diane Williamson. Cambridge University Press, 2018.

Journal of Ethics and Social Philosophy
vol. 30, no. 6 · october 2025
https://doi.org/10.26556/jesp.v30i6.3480

MAKING SENSE OF POLITICAL NORMATIVITY

Mike Gadomski

Political life, like the rest of life, makes demands on us. There are certain things we ought to do and certain evaluations we ought to make when it comes to politics. Some of this normative action is of the prudential variety, some of it is epistemic, and much of it is blandly instrumental. But it seems like a good chunk of it—the chunk that political philosophy cares about—is moral.[1] It is morally wrong, for example, for a government to oppress its citizens. And we have moral reasons to pursue political projects that will make society more just. When political philosophers theorize about justice, legitimacy, democracy, freedom, equality, and so on, the goal is to figure out what morality says about these issues. Or so the standard story goes. Call this *political moralism*.[2]

Political realists find this story problematic. They critique the primacy of the moral in political theory.[3] The traditional way to do this is to highlight the

1 However, see the recent interest in the field of political epistemology, as exemplified in, e.g., Hannon and de Ridder, eds., *The Routledge Handbook of Political Epistemology*.

2 The term 'political moralism' comes from Bernard Williams. He uses it to refer to views that, according to him, "make the moral prior to the political" ("Realism and Moralism in Political Theory," 2). Since Williams, there has been some debate about how to understand moralism. In their well-known 2018 paper, Jonathan Leader Maynard and Alex Worsnip characterize the view as holding that "what gives political theory its normativity is … morality" ("Is There a Distinctively Political Normativity?" 765). In other words, it denies the existence of a distinctly political domain of normativity. More recently, however, Leader Maynard has offered a less ambitious definition of moralism, according to which "moral reasoning is a necessary, though not sufficient, task of political theory" ("What Is Political Moralism?" 14). Leader Maynard still denies that there is a distinctly political domain of normativity; the purpose of the weaker definition is to allow for other kinds of nonmoral (e.g., prudential) reasons to play a role in politics. Because this article is about whether a distinctly political domain of normativity exists, I follow the earlier usage of the term 'moralism'. As I clarify in section 5 below, I am open to the weaker version of moralism.

3 Williams describes this as the "refusal of a mere moral normativity" ("Realism and Moralism," 10). The literature related to political realism's revival is vast and growing. Williams's posthumously published collection *In the Beginning Was the Deed*, particularly its first essay, is a crucial touchstone. For a more radical realism, the work of Raymond Geuss is influential. See especially Geuss, *Philosophy and Real Politics*. A useful critical overview can be found in

importance of instrumental or prudential concerns.[4] Another way is to add to morality's list of competitors by claiming that there is a distinct *political* domain of normativity.[5] Suitably distinguished from the former, the latter strategy has not been particularly popular—indeed, many realists have distanced themselves from it.[6] This is in part because it is not immediately clear what it means to say that there is a distinctively political normativity, nor what it would take to vindicate such a claim. Thus, realists might be better advised to go with the other options just mentioned or to otherwise try to articulate their concerns within a picture that still gives pride of place to morality.

Instead of these more concessive courses, this article argues that there is a readily available route to the idea of a distinctly political kind of normativity. An important early stop on this route is the thought that a picture of normativity that admits of different normative domains is an effective way to make sense of normative conflicts—that is, situations where normativity pulls us in different directions. This suggests a strategy for individuating domains of normativity: showing that the existence of distinct domains is what best explains some particular normative conflict. More specifically for our purposes, if there are conflicts for which a distinctly *political* domain figures in the best explanation, then we have good evidence that there is such a domain. Showing that there are such conflicts, then, is at the heart of the argument. This naturally leads to other questions about the nature of the political domain. I have plenty to say about this as well. But I want to stress that even if one is not convinced by that account, one can and, I believe, should still accept that the question that the debate turns on is whether a political domain, however it is cashed out, is needed to make sense of the normative landscape. In other words, the substantive account of the political domain of normativity that emerges is one contribution of this article. Another, no less important contribution is laying out the strategy for getting there.

The article is organized as follows. In section 1, I introduce the idea of normative domains and how they help us make sense of normative conflicts. In

Rossi and Sleat, "Realism in Normative Political Theory." More recent collections of note include Sleat, ed., *Politics Recovered*; and Sagar and Sabl, eds., *Realism in Political Theory*.

4 See, for example, realism in international relations. For a contemporary effort to cash out the idea of the political's distinctive normativity in terms of the instrumental and the epistemic, see Burelli and Destri, "The Sources of Political Normativity." See also Rossi, "What Can Epistemic Normativity Tell Us About Politics?"

5 See the numerous citations in Leader Maynard and Worsnip, "Is There a Distinctively Political Normativity?" 757n3, to which we can add Jubb, "On What a Distinctively Political Normativity Is"; and Burelli, "Political Normativity and the Functional Autonomy of Politics." See also the recent special collection on political normativity in *Topoi* (2024).

6 Sleat, "Realism and Political Normativity."

section 2, I make the initial case that there is a familiar tension in politics and political philosophy for which political normativity can serve as a good explanation. Section 3 bolsters this case by interpreting the domain of political normativity in terms of the value of a political order making sense to its subjects as authoritative (which I will abbreviate, following Bernard Williams, as 'MS'). With the basics of the theory on the table, I then turn to its clarification and defense against moralism. Section 4 responds to the moralist objection that MS, if it is valuable, is only morally valuable. In section 5, I address another moralist objection, that political normativity is either not normative at all or normative only in the weak and uninteresting sense that etiquette is normative. Section 6 concludes.

1. DOMAINS OF NORMATIVITY

Normativity speaks in many voices. As Derek Baker writes,

> What ought I to do? A lot of things, and frankly, too many things. Too many things, because there are too many oughts. Morally I ought to give to charity, prudentially I ought to invest. Epistemic reasons may demand that I begin to doubt my friend's innocence, while loyalty forbids it.[7]

He gives other examples: the director Zack Snyder has reasons of self-interest to sign on to projects like *Batman v Superman*, though he has aesthetic reasons not to; Gyges had prudential reasons to kill the king, but moral reasons to refrain.[8] We can come up with more: Paul Gauguin's *Two Tahitian Women* is aesthetically valuable yet morally fraught, given the artist's personal history and that history's connection to the content of the painting.[9]

7 Baker, "Skepticism About Ought *Simpliciter*," 230 (emphasis removed). For more on the epistemic conflicting with the moral, see Fricker, *Epistemic Injustice*, 126.

8 Baker, "Skepticism About Ought *Simpliciter*," 234.

9 In an essay about political realism, Matt Sleat also gives the example of aesthetic and moral concerns pulling in different directions when judging a piece of art. Sleat, "Politics Recovered," 6. The example of Gauguin has provided good fodder for those interested in this phenomenon (although it is worth noting that Williams raises Gauguin's case to make an entirely different point about moral luck). Gauguin deserted his wife and children to live in French Polynesia, where he engaged in numerous abusive relationships with young women. These problematic escapades produced notable works, including *Two Tahitian Women*. The artist Michelle Hartney's "performance/hang and run/call to action" *Separate the Art from the Artist* is one example of artists themselves grappling with Gauguin and others (Picasso and Balthus, in Hartney's case) and the issues that they raise. For more on the relationship between the aesthetic and moral (and an early statement of the idea of encroachment), see Eaton, "Integrating the Aesthetic and the Moral." For an excellent

In describing these situations in this way, we make use of the idea of normative *kinds* or *domains*, which I define here, following Errol Lord, as systems of evaluations, recommendations, and demands.[10] The idea of normative domains is relatively uncontroversial, even if there is debate about how exactly to understand them or what ought to count as one.[11] The domains invoked above—the moral, the prudential, the epistemic, and the aesthetic—are among the paradigmatic domains. And the distinction invoked when claiming that political normativity is a distinct kind of normativity is the same distinction invoked when theorists talk about there being moral normativity, prudential normativity, epistemic normativity, and aesthetic normativity.[12]

In addition to illustrating the notion of the domains, the above examples also point to why domains are so useful: they allow us to make sense of the variety of normativity, in particular the ways in which we are so often pulled in different directions. This is not to say that the pulls have equal force; in many cases, it may be that one of the domains wins the day, and sometimes this will be obvious.[13] But even in such cases, domains allow us to account for the remainders, to register the fact that when we hang the Gauguin painting on account of its aesthetic value, we do so at some cost, even if it is the case

treatment of the issue outside of the philosophical literature, which sees things differently from Hartney, see Nelson, "Art Song," esp. 21–23.

10 See Lord, "Choosing the Right Companion," sec. 2.1. We should not take the term 'system' too seriously here, as domains can be relatively disorganized.

11 For similar lists, see Lord, "Choosing the Right Companion," sec. 2.1; Leader Maynard and Worsnip, "Is There a Distinctively Political Normativity?" 756; and Baker, "Skepticism About Ought *Simpliciter*," 230–52. There is some controversy here, particularly around the prudential and its relationship to the moral. Indeed, one group of people for whom talk of domains would likely be rather bizarre is Ancient Greek philosophers. Julia Annas, for example, argues that Ancient Greek ethical theory does not give a role to distinctively prudential reasoning ("Prudence and Morality in Ancient and Modern Ethics"). In the modern context, Alex Worsnip has recently argued that there are no prudential reasons ("Eliminating Prudential Reasons"). For a famous argument as to why we should not collapse prudential reasons into morality, see Wolf, "Moral Saints." For a recent discussion of Wolf's argument, see Lord, "Impartiality, Eudaimonic Encroachment, and the Boundaries of Morality." For a recent monograph on prudential normativity, see Fletcher, *Dear Prudence*.

12 Indeed, this is explicitly how Leader Maynard and Worsnip frame their paper "A Distinctively Political Normativity?"

13 Baker, though, argues against this thought, which would involve an overriding or all-things-considered ought (the *ought simpliciter*, as he calls it). Nothing in my argument hangs on this question. All I am pointing to is the existence and usefulness of domains, which is presupposed by the argument about the ought simpliciter. For discussion, see Musso, *The Historical and Contemporary Significance of Anscombe and Foot's Metaethical Thought*, esp. ch. 4; Tiffany, "Deflationary Normative Pluralism"; and Copp, "The Ring of Gyges."

that hanging it is indeed what we should do, all things considered.[14] Without domains, our normative thought and talk would be flattened and would fail to do justice to normativity's richness.

With this in mind, debates about domains or kinds of normativity can be usefully understood as debates about how best to explain the evaluations that are apt or the reasons that we have. That is, we can read cases like the above as inferences to the best explanation wherein the explanandum is a normative conflict, and the explanans is the existence of distinct normative domains. Understood this way, these situations of conflict are good evidence of there being different domains of normativity. To take one of the cases from above, the distinction between the aesthetic and the prudential best explains the conflictual normative situation Zack Snyder is in when he is asked to direct *Batman v Superman*.

The idea that distinct domains are explanatorily powerful is not novel. It has proven useful in discussions of supererogation, for example. To briefly explain: most of us think that while it might be morally best to put oneself at great risk to save another, it is not required. Actions like these, which go beyond the call of duty, are said to be supererogatory. Supererogation raises a puzzle for moral theory: if one action is better than another, why is the better action not simply what morality demands? That is, if it really is morally better to run into the burning building to save someone's life, why is this not simply what we ought to do? A leading contemporary candidate for a solution to this puzzle is to introduce a nonmoral domain of normativity: while morality does indeed demand the further action, that action is not all-things-considered obligatory because there are normative considerations of another kind that weigh against it. As Daniel Muñoz puts it, "heroic sacrifices are optional, even though they are morally better, because we have mighty non-moral reasons not to harm ourselves."[15] In other words, the initial analysis of the situation appears paradoxical because it makes use of only a single domain. The introduction of a nonmoral normative domain avoids the puzzle and best explains the data.[16] This is akin to

14 As Lord puts it, we want "to be able to register a certain sort of complaint," even when the overall right thing is done ("Impartiality, Eudaimonic Encroachment, and the Boundaries of Morality," 121). The notion of domains helps us to do this.

15 Muñoz, "Three Paradoxes of Supererogation," 702.

16 Muñoz summarizes: "We need two dimensions to make sense of supererogation" ("Three Paradoxes of Supererogation," 702). As Muñoz points out, this does involve a slight revision in the concept of supererogation; going beyond the call of duty is not actually better overall, it turns out. But importantly, it holds onto the strong intuition that it is morally best. Another thing worth noting is that the term 'dimensions' is ambiguous in a way. Indeed, Muñoz is happy to grant that there are other ways of construing the distinction than relying on the idea of a distinctive normativity. But these other ways rely on

our analysis of the other cases involving Snyder, Gauguin, and so on. One need not agree with these analyses in every case; the point is to illustrate that for a nontrivial array of normative situations, we cannot make sense of them from only the moral domain. Indeed, no single domain can do all of the explanatory work by itself in such cases.

This idea of distinct normative domains, as well as the notion that they are explanatorily useful, is very important for what follows. Our strategy moving forward is to identify cases like the above in political contexts, where the existence of a distinctly political domain is the best explanation of the situation. Our analyses of such cases are structurally analogous to the nonpolitical examples given in this section.

2. Normative Conflicts in Politics

To see the sorts of political cases I am interested in, consider the following examples.[17]

> *Immoral Democracy*: A democracy passes, with wide support, a law instituting a system of taxation that is deeply inconsistent with whatever the true theory of distributive justice is.

> *Moral Autocracy*: Disregarding the wishes of the people, who much prefer an unjust tax code, an autocrat institutes the tax code most consistent with the true theory of distributive justice.

How should we think about these cases? On the one hand, Immoral Democracy merits our moral disapproval. The law is unjust, and the people who support it are morally wrong in doing so. On the other hand, something has gone right. (We can remain vague for now about what that is.) The law has a certain kind of normative standing that it would not have if it were instituted in the face of wide opposition. It deserves a certain kind of respect, or some sort of positive evaluation. There are, then, conflicting judgments here. It is important that in pointing this out, we set aside the question of the relative weights of these judgments, as well as the question of what the correct all-things-considered judgment is. The point here is simply to notice the conflict. Similarly,

collapsing the distinction between morality and prudence, a distinction that I think we have good reasons to maintain.

17 Like the examples from section 1 above, these are toy models in that they omit many details that we would expect in the real world. Their purpose is to help us distinguish different normative considerations; given this purpose, their schematic nature is warranted at this stage of the argument.

Moral Autocracy yields competing evaluations. Despite the fact that the morally superior law has won the day, something has gone wrong here. At the very least, there is a troubling distance between the people and the political power they are being made to comply with. Again, setting aside the question of the all-thing-considered evaluation, there is a normative conflict here.

These thoughts need sharpening, and we will get to that in a moment. For now, note that such cases bear an important resemblance to those considered in section 1 above. Take, for example, a simple case of supererogation. Suppose that I have a fair amount of disposable income, and consider two options regarding how I could dispose of it: give it to those worse off than me or spend it on various personal pursuits and pleasures such as travel, art, restaurants, gifts for loved ones, and so on. Morality tells me that the first option is best. The second option, or at least certain versions of it, might be morally permissible, but it is morally inferior to the first. On the picture we have worked up so far, the best way to analyze this is by holding that my reasons to choose the first option are moral, and my reasons to choose the second are not. Whatever the all-things-considered verdict might be, figuring it out involves weighing these different kinds of reasons.

Compare this with the situation from our two political cases in which the governments decide on new tax codes. Tax code *A* is most just, but everybody prefers tax code *B*. Which code ought the government implement? They have moral reasons to implement tax code *A*. But they also have reasons—of some kind—to implement tax code *B*, given that this is what the people prefer. Again, whatever the all-things-considered verdict might be, figuring it out involves weighing these reasons.

A natural way to cash out the supererogation case is to say that we have moral reasons to do what morality demands and reasons of another kind—prudential, or perhaps eudaimonic—to pursue our personal projects.[18] Similarly, I argue, a natural way to cash out the political case is to say that we have moral reasons to pursue political projects that morality demands and reasons of another kind—which I call *political*—to pursue political projects that align with our own values and commitments. My argument is that in both cases, there is a normative conflict that morality alone cannot explain. And in both cases, a nonmoral normative domain can be invoked to fill in the explanatory gap. Again, it may or may not be the case that at the end of the day, what we really ought to do, all things considered, is what morality demands. This (controversial) idea is beside the point.[19] The point is that there are two domains of

18 Lord, "Impartiality, Eudaimonic Encroachment, and the Boundaries of Morality."

19 See note 13 above.

normativity—two systems of evaluations, recommendations, and demands—that exert themselves in such situations. Whatever the all-things-considered verdict might be, figuring it out involves considering these different domains.

The core argumentative strategy for establishing the existence of a distinct political domain of normativity is now up and running. Describing this strategy is itself one contribution of this article. In what follows, I am concerned with the other contribution: pursuing that strategy to its end. I have sketched the initial argument for doing so in this section, but more work needs to be done. Two tasks in particular loom large. For one, we must give some idea of what the political domain is, ideally in a way that resembles the way in which we have a grasp on what the paradigmatic domains are and in a way that adds plausibility to the explanation of these cases. Second, we must show that alternative explanations that do not invoke a nonmoral domain are inadequate. I address these tasks in sections 3 and 4, respectively.[20]

3. Political Value

First, let us say more about what this distinctly political domain of normativity is all about. Note that this is a distinct question from the question of what politics as a domain of human activity is all about. Many realist approaches to political normativity try to define politics as an activity in terms of some distinguishing characteristics and then to build up a theory of the kind of normativity exclusive to that domain.[21] This strategy faces certain difficulties, however. For one, it is not easy to define politics in a succinct and uncontroversial way. More importantly for our purposes, reflection on the paradigmatic normative domains reveals that they are not solely contained within particular spheres of human activity. The epistemic, for example, is at the heart of the activity of inquiry, but it also plays a role in many other areas of human life. Similarly, the aesthetic domain is at the heart of artistic activity, but it guides our lives in other ways as well. This suggests that we should not start from a characteriza-

20 Once we have the idea of the distinct political domain, other questions arise, especially having to do with that domain's significance. I address this issue in section 5 below.

21 For example, one may understand politics as the domain of activity concerned with enforcing conduct amidst moral disagreement or the domain concerned with ensuring social order. For discussion, see Leader Maynard and Worsnip, "Is There a Distinctively Political Normativity?" 767–73, 781–85. To take another example, in his investigation of the nature of political value, Sleat argues that political values must be consistent with the nature of political practice ("What Is a Political Value?"). He seems, then, to be thinking of the political as a domain of human *activity* that other values can be appropriate or inappropriate for, not a domain of *normativity* in the sense that the moral, the aesthetic, and so on are domains of normativity.

tion of political activity from which we try to theorize the kind of normativity exclusive to that domain. In working up a theory of political normativity, then, I avoid definitional claims about what politics is as a domain of activity.[22]

Instead, a helpful alternative suggestion for proceeding comes from Lord, who points out that we can profitably understand many domains of normativity in terms of their central organizing values.[23] For example, beauty and truth play this role in the domains of the aesthetic and the epistemic, respectively.[24] Articulating such a value for the political domain would go a long way toward filling out the argument that such a domain exists and is distinct from morality.[25]

It also helps to have a rough characterization of the moral domain in hand or, at the very least, to avoid any controversial commitments about it.[26] If it turns out, for example, that the distinctness of the political rests on a contentiously narrow conception of the moral, this would make the resulting view much less interesting. Indeed, as Leader Maynard points out, this has sometimes happened in the realism-moralism debate.[27] We want to avoid stacking the deck in the realist's favor in this way. Using the idea of a central organizing value, then,

22 This has the happy consequences of avoiding certain difficulties inherent in views that employ the activity-based strategy. Leader Maynard and Worsnip, for example, criticize views that foreground disagreement and the priority of legitimacy ("Is There a Distinctively Political Normativity?" 767–73, 781–85). Eva Erman and Nicklas Möller characterize an entire family of realist views as holding that political reasons are just those reasons that are operative in politics (where politics is understood as a domain of activity rather than as a domain of normativity) ("The Role of Moral Norms in Political Theory"). As they point out, such views have been roundly criticized.

23 Lord, "Choosing the Right Companion." There is more to say about the idea of a central organizing value. For Lord, normative domains are often (but not always) *constitutively* organized around such values, which means that there is some attitude that is constitutively tied to the value, as belief is tied to truth (this is to say that belief has an essential connection to truth; in this case, belief's standard of correctness is tied to truth). A full account of political value might include a discussion of these issues, but I leave such details out here. For now, my ambition is simply to get an idea of political normativity off the ground. For a helpful related discussion regarding the moral's relationship to the epistemic, see Julia Driver on intellectual and ethical virtues in Driver, "The Conflation of Moral and Epistemic Virtue," as cited in Fricker, *Epistemic Injustice*, 126.

24 The aesthetic case, as Lord notes, is a bit complicated, as there are significant debates about beauty's relationship to the aesthetic ("Choosing the Right Companion," sec. 5.2). Nothing for either of our purposes hangs on these debates, though.

25 That said, we might be compelled by the argument for a distinctively political normativity laid out in sections 1 and 2 and yet disagree on the political's organizing value.

26 I thank a referee at this journal for pushing me to address this issue.

27 Leader Maynard, "What Is Political Moralism?" 14–16. As a referee also points out, this is not surprising, given political realism's connection to Williams, whose well-known conception of morality is too restrictive for many theorists.

we can posit *giving due regard to others* as having this role for the moral domain. This particular locution is derived from Lord, but the underlying idea is not idiosyncratic.[28] Leader Maynard, for example, argues that *other-regardingness* is the conceptual core of the moral domain.[29] This notion has the advantage of being both intuitively plausible and capacious enough to include many different ways of cashing out exactly what it means to give due regard to others.[30] It includes, for example, consequentialist views that center on aggregate utility or well-being; on such views, regarding others properly involves considering them as contributors to this aggregate.[31] Contrast this with, say, the Kantian notion of regarding others as ends in themselves. Such views can all be understood as moral in that they are essentially interested in the importance of giving due regard to others.[32]

Returning to the political, then, what can play the requisite role? The cases from the previous section—Immoral Democracy and Unjust Autocracy—as well as existing work in political realism give us some good starting points. Drawing on the work of Williams in particular, I suggest that we understand the value that competes with morality in such cases to be the value of a political order that is aligned with its participants, particularly with their desires, needs, aspirations, and so on. But we can specify this further. As Williams emphasizes, it is a basic and universal fact about human societies that they continually ask and attempt to answer questions about why power in their society is arranged as it is.[33] This practice points to a crucial goal of politics: to achieve a situation in

28 Lord, "Choosing the Right Companion," sec. 5.3.

29 Leader Maynard, "What Is Political Moralism?" 15.

30 Leader Maynard, "What Is Political Moralism?" 14–16.

31 Opponents of consequentialism of course deny that this counts as giving people due regard—indeed, one of the most famous objections to consequentialism is precisely that in subordinating individuals to the aggregate, it does not give them due regard—but we can understand this as a debate about what it means to give due regard to others, not about whether morality is about giving due regard to others. For consequentialists, as I understand them, there is no notion of giving due regard to others outside of considering them as vessels of value in a sea of aggregate value. Some consequentialists may want to add that we can and should also take up another, more personal point of view of others, as Peter Railton famously argues ("Alienation, Consequentialism, and the Demands of Morality"). Again, though, such two-level or "sophisticated" views are just more examples of how to think about giving due regard to others. I thank a referee at this journal for encouraging me to think through these issues more.

32 The capaciousness of this notion of giving due regard raises the suspicion that it must include whatever values are at the heart of political normativity. This is an important concern; I say something about it momentarily and address it more fully in section 4 below.

33 Williams, "Realism and Moralism." This is a point also at the heart of Thomas Piketty's *Capital and Ideology*, specifically as it relates to inequality. See also Greene, "Legitimacy

which people—as many as possible and to as great of a degree as possible—see the various forms of power under which they live as making genuine claims on them. Williams refers to this as the order *making sense* (MS, as I will follow him in abbreviating) to its subjects as authoritative.[34] It is this notion, I argue, that we should understand as at the heart of the political domain of normativity, playing the role that beauty, truth, and giving due regard to others play for the domains of the aesthetic, the epistemic, and the moral, respectively.

It may initially appear that MS is a moral idea, especially if we understand morality as being about giving due regard to others. I more fully address this concern in the next section. For now, though, two points are worth making. The first is that MS is not at its core an other-regarding notion—that is, its value is not reducible to the value of people being treated or considered in a certain way.[35] Its value just is what I have said above: that of being able to make sense of one's political order in a particular way—namely, in that one can affirm it as being genuinely authoritative.[36] My claim is that such a situation is good—and good in a significant way—independently of any relational or other-regarding considerations, just as beauty and truth are. Put another way, the value at the

Without Liberalism," 315; and Greene, "Is Political Legitimacy Worth Promoting?" 88.

34 Williams, "Realism and Moralism," 10–11. The *as authoritative* part is crucial. Lots of power structures can make sense, but not all make sense *as* the sort of structure that makes a claim worthy of heeding.

35 It does, however, have implications for how to treat others. But this does not make it moral, as I explain in section 4 below.

36 A full theory of political normativity would say more about the question of what kinds of orders tend to MS and why they do so. Space prohibits developing such a theory here, but a few initial thoughts are worth giving. For one, the answers given to questions about MS vary depending on particular historical circumstances. Second, as I indicate below, democracies tend to be a reliable vehicle for MS, though this is not a conceptual truth. Third, following Williams, we might think that one consistent truism is that might does not make right—i.e., political orders that justify themselves only in terms of raw power tend not to make sense as authoritative. Fourth, power must solve what Williams calls the "first" political question, which is "the securing of order, protection, safety, trust, and the conditions of cooperation" (Williams, "Realism and Moralism," 3). This is not sufficient, though: a state, for example, may answer this question in a way that does not MS to its subjects. One possible suggestion to generalize this notion is that political orders in general make sense when they solve the problems their subjects take themselves to have. Javier Rodríguez-Alcázar, for example, argues that politics should be understood as a kind of collective instrumental rationality that aims at achieving the goals of the people, whatever those goals may be ("Beyond Realism and Moralism"). Similarly, Vitor Sommavilla proposes that political normativity is closely related to the solving of political problems ("A Constructivist Account of Political Normativity"). I am friendly to both these ideas, though I understand the connection between achieving goals (or solving problems) and MS as contingent (though robust) rather than analytic.

heart of political normativity is the value of being part of a political order that one can affirm, not the value of people being given due regard.[37]

Second, we should resist the temptation to understand MS's value solely in terms of its apparent instrumental connection to the moral goods of peace and stability. On the one hand, it is true that societies that do poorly with regard to MS tend not to last very long or must undergo significant transformation.[38] And it is true that governments might have instrumental moral reasons, grounded in the moral demand to avoid violence, to achieve MS. But the demands of MS can diverge from the demands of peace and stability. Indeed, the fact that people are willing to unsettle their own political orders when they do not make sense is evidence that they value MS not merely in virtue of its connection to peace and stability. So long as we take such actions seriously, we take seriously the value of MS for its own sake.[39]

With that said, then, we can highlight the significant explanatory power of MS. Most immediately, it has explanatory power in the particular way that interests us for this argument—namely, it explains the kind of normative conflict that we considered in the previous section through the two main examples. Furthermore, MS is a familiar notion that figures into a great deal of political philosophy, especially in the liberal tradition. It is related not only to influential discussions of democracy, but also to important discussions of the significance of consent, legitimacy, and self-determination, to name a few.[40] And as I have

37 This is compatible with thinking that in some cases, part of what it is to give due regard to others is to subject them to a political order that makes sense to them (or at least to not subject them to a political order that does not make sense to them). I discuss this more near the end of section 4 below.

38 As Piketty documents exhaustively in *Capital and Ideology*.

39 Interestingly, this distances my view from one way of thinking about political realism as being based on the priority of order and stability over other ends. For example, Erman and Möller have recently characterized realism in this way ("The Role of Moral Norms in Political Theory," 3). In this sense, I am friendly to Samuel Bagg's enjoinment to disconnect the realist project from the priority of legitimacy ("Realism Against Legitimacy," 29–60). I am grateful to a referee at this journal for suggesting that I address the issues raised in these two paragraphs.

40 For the significance of consent, see the entire social contract tradition. For a well-known example connecting MS to legitimacy, recall Michael Walzer's claim that "a state is legitimate or not depending upon the 'fit' of government and community, that is, the degree to which the government actually represents the political life of its people" ("The Moral Standing of States," 214). Or see public reason accounts of legitimacy (in, e.g., Quong, "Public Reason"), according to which political principles must be justified to those subject to them. (We return to public reason below.) Finally, for an example connecting MS to self-determination, see David Miller, who argues that the intrinsic value of self-determination is "the value of belonging to a group that can act so as to make a difference to the world

mentioned, it has sociological significance: people and societies throughout history have tried to achieve it. While this latter point does not immediately vindicate MS as a reputable notion from the point of view of normative theory, it does lend it some legitimacy.[41] It also partly explains its interest to political realists, who are characteristically interested in bringing political theory closer to the real world of politics.

In any case, the idea of making sense in this way is not merely sociological. It figures into a plausible reading of the cases from the previous section. For example, when we assess a democracy that has run afoul of justice (as in Immoral Democracy), a good way to understand the recognizably positive aspect of the situation is in terms of MS. The law's democratic credentials suggest (though may not guarantee) that it reflects the values and commitments of the society's members; it suggests, in other words, that the use of power in this way, for these ends, makes sense to the society's members, given who they take themselves to be and the problems they take themselves to have.[42] Likewise, when considering the situation from the point of view of the citizens, we can put some more meat on the bones of the thought that they have reasons to want the government to pass tax code *B* (the one that they prefer) instead of tax code *A* (the just option). Those reasons are *political* reasons, grounded in the value of achieving MS, the situation wherein the power under which they live makes sense to them as authoritative. The enactment of tax code *A* would take them further from MS, even if it would take them closer to morality.[43]

We have now, therefore, an idea of a political domain of normativity, organized around the central value of MS and motivated by its explanatory power in political conflicts. There are some difficult and important questions to be raised

---

in accordance with the formed will of its members" (*Is Self-Determination a Dangerous Illusion?* 36).

41 It is worth recalling John Stuart Mill's famous point that good evidence for something being valuable is that people do in fact value it (Mill, *Utilitarianism*, 35). This is also related to a background methodological commitment of mine (and Williams's) of pursuing philosophy as a humanistic rather than scientistic discipline: the primary goal is to understand ourselves and our activities. See Williams, "Philosophy as a Humanistic Discipline."

42 A referee points out that in more realistic versions of this case, we may have reason to doubt that such a society really does achieve MS. If a government taxes its citizens in ways that, for example, privilege certain groups over others, it seems unlikely that they are ruling in a way that makes sense to its subjects as authoritative. The first point to reiterate is just that I stipulate for the sake of argument that MS is indeed achieved, so as to illustrate the distinction I am drawing. But I am also not sure that this is so far from some realistic cases. Many unjust tax codes are widely supported, or at least seem to be so.

43 Again, I make no claims about what they should want all things considered; I am arguing only that they have political reasons to want tax code *B* and moral reasons to want tax code *A*.

about this picture, and they occupy the remaining substantive sections of this article. Before turning to them, though, I want to motivate this idea further by pointing to two areas in which the idea of political normativity, understood in terms of MS, has helpful theoretical potential. The point here is not to settle any debates in these areas—this would be far too ambitious in the current context. My goal is modest: I mean only to illustrate the promise of political normativity for helping with some problems that have characteristically beset long-running debates.

The first area is democratic theory. MS gives us a way of understanding what is good about democracy, and it does so in a way that has important advantages over the traditional ways in which democratic theory has tried to do so. Such views are usually divided into two categories.[44] On the one hand, instrumentalists ground democracy's value in its tendency to deliver morally good results.[45] They may consider cases where democracy acts wrongly—cases like Immoral Democracy—to be unfortunate exceptions to this general rule, but they must concede that in those instances, democracy is not in fact good. Many theorists, though, prefer a commitment to democracy that is not so contingent, that is not held hostage to the results turning out in democracy's favor.[46] So they opt for an intrinsic value account, according to which democracy is itself part of what justice demands.[47] These theorists, though, find it difficult to explain why democracy is intrinsically good when it is so easily compatible with injustice.[48] Democracies, after all, can (and do) elect racists and fascists.

The political domain, as I am sketching it, understands democracy's value in connection to its ability to achieve MS. Citizens of a well-functioning democracy have a plausible story to tell each other about why they should accept how things have turned out (that is, about why their political order MS): it is what they have chosen, together, even if not everybody agrees. Or, if they are not moved by considerations regarding their own agency and political participation,

44 For a good introduction, see Christiano and Bajaj, "Democracy."

45 See, for example, Arneson, "Democracy Is Not Intrinsically Just." There are different ways to cash this out, but all are united by the thought that democracy's value is to be found in its outcomes, particularly in comparison to alternatives.

46 Many also want to avoid saying that, as Valentini puts it, "a society governed by a wise sovereign, or a small enlightened elite, is fully just, even if it implements an equitable distribution of resources" ("Justice, Disagreement, and Democracy," 181). In other words, they want to avoid being committed to even the possibility of a just autocracy, at least in practice.

47 Various theorists have argued, for example, that democracy is the only form of government compatible with human freedom or human equality. See Christiano and Bajaj, "Democracy," especially secs. 2.2 and 3.2.

48 Valentini, "Justice, Disagreement, and Democracy," 180–82.

they may point to the fact that, as Samuel Bagg argues, democratic institutions are reliable if limited bulwarks against state capture by self-interested elites.[49] In either case, these kinds of commitments to democracy are more robust than other instrumentalist arguments that rely on contingent outcomes, yet more realistic and more plausible than intrinsic value theories that try to accommodate democracy within morality and thus have difficulty with the fact that democracies can and do make immoral decisions.[50]

Furthermore, unlike intrinsic value theories, this view also explains the fact that one of the core goods that democracy delivers can also be delivered in nondemocratic contexts. This is just to say that nondemocratic political arrangements can and do sometimes make sense to their subjects and that something valuable is achieved in such cases. But we can allow for this fact while seeing why democracy is still a more reliable route, even if it is not the only route to achieving that value.

The second area of interest is legitimacy, particularly the liberal idea of legitimacy. One of the most influential theories of liberal legitimacy ties it closely to the ideas of public reason and public justification, the basic thought of which is that political power is legitimate only when exercised according to reasons that its subjects can accept.[51] This is seen by its proponents as the best solution to the fundamental question of how to reconcile coercive political power with the freedom and equality of persons. Only by giving reasons that are acceptable to subjects themselves can rulers coerce them in ways consistent with their freedom and equality.

Public reason liberalism is thus based on an ideal that looks very similar to MS—what David Enoch calls "justification-to."[52] Enoch raises a number of problems for this notion. One is closely analogous to the problem that we have already discussed, faced by intrinsic value theorists of democracy. This is the problem arising from the fact that people can be mistaken in their beliefs. Just as democracies can pursue unjust projects, people can find abhorrent regimes acceptable to them (or, conversely, they can fail to find just regimes acceptable).

49 Bagg, "The Power of the Multitude," *The Dispersion of Power*.

50 As a referee points out, the argument thus far still leaves available a view on which MS's value is both intrinsic and moral but also *prima facie* (in W. D. Ross's sense). Such a view partly captures the intrinsic value of democracy while also explaining the existence of immoral democracies, without needing a nonmoral domain of normativity. I address such a view in the next section.

51 For the Rawlsian formulation of this idea, which has some important qualifications, see Rawls, *Political Liberalism*, 217. See also Quong, "Public Reason."

52 As Enoch puts it, "for a state (for instance) to be legitimate, its authority must be justifiable to each of those subject to it" ("Against Public Reason," 115).

This means that public reason theory needs to introduce significant idealization, which it does with the idea of the "reasonable."[53] But Enoch argues that this idealization is ad hoc—if it is important to give reasons to reasonable people that they can accept, then it is important to give reasons to *everyone* that they can accept, including the unreasonable.[54] In other words, the idea of the reasonable seems rigged up just to escape the counterexamples where people accept injustice.[55]

This nicely throws into relief a general problem that political philosophy faces when trying to fit concerns related to MS entirely into the moral domain: it can do so only in ad hoc ways because of the persistent threat of conflict between what morality demands and what makes sense as authoritative to people. Enoch thinks that the solution is to downgrade the importance of justification-to (or MS) and to consider it just as one moral desideratum among others.[56] In the next section, I argue that this solution fails. On my view, political liberalism is right to emphasize justification-to, but Enoch is also right that it does so in an unconvincing way. My diagnosis is that this stems from political liberalism's failure to properly distinguish the moral and the political, not from its being too concerned with justification-to.

Before coming to that, though, consider another of Enoch's criticisms. He claims that the importance that public reason theorists place on justification-to stems from a fundamental confusion.[57] The confusion is this: when a ruler coerces a subject, the reasons they do so are just the things they would cite if they were to give an honest explanation of their actions. Their reasons are not the fact that they believe those things. On this, Enoch gives the following argument:

> Let's imagine a possible world in which you are mistaken about autonomy, thinking that it's not of value, even though it is. Do you still want—in the actual world—to impose autonomy-based directives on people in that hypothetical world? A positive answer seems obvious. But in that world, you don't believe in the value of autonomy. Still, in that world autonomy is of value. So what we can learn from the fact that when you think (in the actual world) about that possible world you still want to impose autonomy-based directives is that your reason for imposing

53 Again, this is analogous to how defenders of the intrinsic value of democracy need to idealize what counts as democracy.

54 Especially since, as Enoch points out, many people are unreasonable in the technical public-reason sense, including him ("Against Public Reason," 121–22).

55 Enoch, "Against Public Reason," 117–30.

56 Enoch, "Against Public Reason," 138.

57 Enoch, "Against Public Reason," 130–34.

> them is *that autonomy is of value* (which is true there too), not *that you believe that it is of value* (because on that world, you don't).[58]

Enoch's target is the connection that public reason theorists draw between justification-to and equality, but this is not my concern here.[59] I want to point out another issue raised by this example, one that illustrates the importance of political normativity and its place in our political thought.

The problem with Enoch's argument is that the answer that he claims is obvious is far from so. From the perspective of morality, we can see his point; if we think autonomy-based directives are what morality calls for, we will want them imposed in both worlds. But there is another important sense in which it would be very strange to want your alternate-world self to impose autonomy-based directives: your alternate-world self does not value autonomy and thus does not believe that autonomy-based directives should be imposed. For them to do so, then, would bring them further away from their ideal of what a society should be like. *This would be a bad thing for them*, in precisely the way that the notion of MS and the domain of political normativity explain: it creates a situation wherein they cannot make sense of their political order as authoritative. Morally, then, we may want our alternate-world self to do as Enoch says, but *politically* we recognize that the resulting political order would be defective in an important way and that our alternate-world self would be alienated from it.[60]

To summarize this section, I have argued that we should understand the political domain of normativity as being organized around the central value of a political order making sense to its subjects as authoritative (MS). Following my proposed strategy, the main motivation for this view is its explanatory power regarding conflicts like the ones found in Immoral Democracy and Moral

58 Enoch, "Against Public Reason," 131.

59 The thought is that when a ruler coerces a subject for reasons that the subject cannot accept, the ruler is placing themself above the subject in a certain kind of way, thus not treating the subject as an equal. Enoch thinks that this is not so; because the reason has nothing to do with the ruler themself, no issue of equality is raised. The success or not of this argument is orthogonal to my concerns here.

60 A feature of the example that might make it hard to see this is that it focuses on the point of view of the ruler rather than the subject. From the perspective of the subject, the point is even more vivid. As a companion example, Enoch points out that it would be silly to think that your belief in the moon's gravity explains the tides rather than simply thinking that the moon's gravity explains the tides. Again, as he puts it, the reason is just not about you. While he considers this case to be analogous to the political cases, we can easily see why it is not—it is entirely unrelated to concerns about power making sense to its subjects. No concerns about political normativity are in the picture. Thus, we can accept his analysis of the moon and the tides without accepting his analysis of the justification of political power. See Enoch, "Against Public Reason," 131–32.

Autocracy. I have also given two examples of how political philosophy tries to accommodate something like MS but runs into problems when doing so. I have argued that the idea of a distinctly political normativity has the potential to avoid these problems. In the case of democratic theory, it can navigate between intrinsic and instrumental views of democracy's value. In the case of public reason, it can capture concerns about justification-to without being ad hoc. To reiterate an earlier point, these two examples are just sketches of arguments. They do not settle the debate about the value of democracy, nor do they amount to knockdown considerations against public reason liberalism, an important and longstanding research program with a voluminous literature. What they are meant to do is demonstrate the theoretical promise that the idea of political normativity, understood as being centered around the notion of MS, has. More work needs to be done to take the next steps in this direction.[61]

4. The Distinctive Normativity of Making Sense

It is time now to address two problems that have lurked in the background since I first introduced the general strategy in section 1. The crux of that strategy is that (1) the existence of distinct normative domains helps explain normative conflicts, and (2) the political domain helps explain a familiar class of conflicts in politics between MS and morality. The first issue is why MS is not a notion that can be handled from within the domain of the moral. Why can morality not capture all of the relevant normative action? The second problem comes from a different angle. It grants that MS is a political notion but asks why it is normative at all. We might accept that it exists and that it conflicts with morality, but does it ever matter? Does it ever win? What kind of authority does it have? I take on the first of these issues here and the second in the following section.

Regarding the first problem, the discussion in the previous two sections already puts pressure on the idea that MS is a moral notion. For one, we have seen that there is a mismatch between what political normativity cares about and what morality cares about. The cases of conflict in section 2, as well as their structural similarity to familiar nonpolitical normative conflicts, bring this out. And in section 3, we filled out this picture by clarifying the central organizing value of the political domain and distinguishing it from the central moral notion of giving due regard to others. We have also seen that MS so understood helps make sense of tensions raised by moralist theories of democracy and liberal legitimacy that try to incorporate something similar to MS.

61 I appreciate the suggestions of a referee at this journal that I clarify the scope of my claims in the section.

We can now press the case further. In doing so, we can also more fully address a seemingly viable alternative picture to the one I have painted, according to which the conflicts I highlight are *intra*domain conflicts rather than *inter*domain conflicts.[62] This view may be a natural one for those attracted to the thought that judgments of moral value are *prima facie* rather than all things considered.[63] On such a view, morality itself often pulls us in different directions as it involves a variety of different values and requirements. So why not think that MS is one such demand? While one response is to simply deny that this is the right way to think about morality, it is important to address the concern more directly in order to avoid the objection that my argument relies on a particular conception of the moral domain endorsed just in virtue of its amenability to my conclusions.

In addition to the issues I have already raised, then, the main problem with this moralist position is that it cannot fully capture the sense in which MS is valuable without delivering implausible verdicts about morality. The best way to see this is to examine our cases from section 2.[64]

Consider first Moral Autocracy, in which the government imposes a morally good measure over the opposition of its people. While there may be disagreement about what the correct, all-things-considered judgment is regarding whether the government ought to do this, I take it that (1) many will feel that they ought not, and (2) even those who think that they ought will feel that there is still a loss to be accounted for and a corresponding complaint to be registered on behalf of the people. But the moralist proposal on the table has a great deal of difficulty explaining this. This is often not apparent in real-world cases

62 In other words, such a view accepts that there are conflicts but holds that they are conflicts between moral values; MS is a moral value that competes with others in politics. As we have seen, this appears to be Enoch's view with regard to what he calls justification-to. See also Leader Maynard and Worsnip, "Is There a Distinctively Political Normativity?" 767.

63 This is helpfully pointed out by a referee, using the example of W. D. Ross's theory. See Ross, *The Right and the Good*. When I use the term '*prima facie*', I use it in Ross's sense, although many now understand him to have been pointing to a notion better captured under the concept of *pro tanto*.

64 There is also the worry that if we understand MS as one moral value among many, it will lose too easily in these competitions with other values. This is analogous to the problem mentioned earlier with instrumental defenses of democracy, that they make the democratic commitment too contingent. In that case, the issue is that reasons for valuing democracy are too frequently overridden by the moral benefits of nondemocratic systems. Here we might analogously worry that if MS is understood purely as a moral value, it is hard to see why we should care about it when it conflicts with other moral concerns, like moving toward a just society. As I point out, Enoch is happy to say that we want autonomy-based directives imposed in a society where people do not value autonomy. From the moralist perspective, why should he not be?

because as a matter of political reality, such autocratic actions tend to involve morally dubious methods.[65] Further, actual situations of enlightened rulers like this are few and far between; the moral lines are rarely so starkly drawn. To make the point more vivid, then, it helps both to set aside the question of coercive methods and to ratchet up the moral distance between the measure being imposed and what makes sense to the people it is imposed on. While this takes us afield of real-world scenarios, it helps to clarify the underlying normative issues. Imagine, for example, laws demanding basic equality and nondiscrimination being imposed on a society full of racial supremacists. What *moral* complaint can these people have, even if just *prima facie*? In what sense are they not being given due regard? Or to return to Enoch's example, imagine that measures with immensely autonomy-boosting effects are being imposed on people who do not value autonomy. Again, it is hard to see why there is anything to worry about in such cases, if considered from solely the moral perspective. And yet it does seem like there is something to worry about.[66] On my view, political normativity is what explains this.

Immoral Democracy illustrates the same point. To see this, it is again helpful to raise the stakes and consider cases where a government does something especially abhorrent with the full support of the people. As with the other case, the proposed moralist assessment here is that one moral value is traded off for another; the situation is good in one way (in the achievement of MS), bad in another (in the immorality of the government's actions), and the senses of both good and bad are moral. Perhaps the value of MS is very weak and is outweighed; in any case, it is understood as morally valuable. This approach fails, though, because it is implausible that it is at all morally good to have an MS relationship to an extremely unjust political order. In fact, as we ratchet up the injustice, the MS relationship itself becomes morally *worse*. It is especially morally troubling when people approve of injustice—that is, when abhorrent regimes not just exist but also make sense to their subjects as authoritative.[67] From the purely moral perspective, then, it is very difficult to hang on to the thought that there is anything good about such situations, let alone something good about the MS relationship in particular. Morality seems to lose out here across the board.

65 My view has no trouble granting this. Indeed, what my argument shows is that it tends to be these methods that do the moral work, not, crucially, considerations about MS itself. I am grateful to a referee at this journal for pointing this out and giving me an opportunity to clarify my view in this regard.

66 It is worth repeating: we are setting aside here the question of the all-things-considered judgment.

67 Fabian Wendt makes a similar point about the value of widely accepted laws ("Compromise and the Value of Widely Accepted Laws," 56–58).

Note that this does not rely on any unduly narrow conception of the kind of normative domain that morality is: due regard, on any plausible view of that demand, is not given in such cases. From the point of view of morality, people with abhorrent moral views are not owed, even in a weak or *prima facie* sense, a political order that enforces those views. At the very least, the burden is on moralists to defend a substantive view of morality according to which they are owed this. And yet there is some kind of positive evaluation, however weak, that seems apt here.[68] Again, political normativity can fill this explanatory gap.

The upshot in both cases, then, is that it is difficult for moralists—even those with a pluralist picture of the moral—to accommodate the value of MS while staying within the bounds of a plausible morality.[69] As a final attempt, though, consider a more sophisticated moralist view of MS, according to which the moral value of MS is conditional on some minimal moral criteria being met. Outside of these conditions, the view would hold, there is nothing normative at all about MS. Furthermore, the defender of this view could also hold that MS is nonetheless noninstrumentally valuable. This would accommodate the fact that MS can diverge from other moral concerns (such as distributive justice), that it is nevertheless to be pursued for its own sake, and, crucially, that its value is moral. Brutally repressive theocracies, for example, might make sense to their subjects as authoritative, but this is not in any way valuable. Minimally decent hierarchical societies, though, provided they do things like provide for the basic human rights of their citizens, might achieve MS in the normative sense despite being far from justice.[70]

Let us grant that this represents the most plausible way for moralism to accommodate the importance of MS without admitting a distinct domain of

68 A moralist might insist that this positive evaluation is moral, but again, they face the uphill battle of explaining these extreme cases in a plausible way, i.e., in a way that is not ad hoc and fits with our other considered moral judgments. It is worth recalling here that my overall argument is an inference to the best explanation. So we can press the question to the moralist: Is morality really the *best* explanation here (especially when there is a ready-made alternative in the form of political normativity)?

69 This point seems to be accepted by many theorists. It is one reason why moral principles that have something like MS built into them also build in limits designed to rule out such cases. That is, they build in limits to prevent very morally bad situations from meeting the principles' criteria. We have already seen this in the cases of intrinsic value theories of democracy and public reason theories of legitimacy. The same is true for theorists who attribute intrinsic moral value to self-determination, wide acceptance of laws, or nonliberal legitimacy, For example, Amanda Greene defends legitimacy as having moral value independently of justice ("Legitimacy Without Liberalism," "Is Political Legitimacy Worth Promoting?"). For her, this value is noninstrumental but conditional. I address views like this in the next paragraph.

70 On decent hierarchical societies, see Rawls, *The Law of Peoples*, 71–85.

normativity. One immediate worry is that such a view is still ad hoc in the same way that Enoch argues public reason's use of the reasonable is ad hoc. In other words, how can we set a principled threshold of morality at which MS switches from being morally invaluable to morally valuable, in a way that is consistent with why we care about MS at all? Because realists understand the situation as one in which MS competes with morality, they avoid this question.

More deeply, a realist has the advantage of being able to hang on to a commitment to the value of MS that is independent of morality. This more sophisticated moralist still gives up the claim that in the unjust democracy, for example, there is something good *in the very democraticness of the situation itself*—namely, MS—not merely MS as combined with some other minimal moral criteria. The moralist says that what is valuable here is MS *in conjunction with* some minimal moral criteria (if it obtains). They then give up the thought that MS itself is *pro tanto* good—that, all else being equal, it is always in some sense better for a political order to make sense than for it to not.[71] From the moral perspective, that question hinges on whether the moral criteria are met. If they are not met, then it is worse if the political order makes sense.

The reason this is a problem is that, as I stress throughout, the *pro tanto* goodness of MS *itself* is an important feature of our political thought. It underlies our intuitions about cases like Immoral Democracy and Moral Autocracy, plays a central role in political theory, and influences our political activity. There is something commendable about democratic outcomes even when those outcomes are morally bad, and there is something bad about unpopular rule even when that rule is morally good. Reasonable people can argue about how strong or weak this something is, but the point is that it is there nonetheless. In short, there is an important, nonmoral way that politics can go well even when it goes morally wrong. The moral perspective alone cannot accommodate this thought. My suggestion for saving it is to introduce the distinctly political perspective.[72]

Before concluding this section, there is a related issue to address. Even if we buy the argument just given that the political domain best explains the value of MS and its conflicts with morality, there still seems to be a suspiciously tight relationship between morality and political normativity. What is morally right seems to have something to do with what makes sense to people as authoritative, and what makes sense to people seems to have something to do with what

71 This is precisely the thought that motivates the first moralist view that I address in this section but that I argue fails to comport with a plausible conception of morality.

72 It helps to keep in mind throughout this argument that the sense of good being used is the broad and general sense, unless otherwise specified. It is the same way we use the word when we say that a work of art is aesthetically good but morally bad.

is morally right. But this is no threat to the distinctness of political normativity. It is just an instance of the more general fact that domains can interact in various ways while remaining distinct. This is uncontroversial but important to make explicit. Consider two sorts of interaction. One sort is *extensional overlap*: the same action can be required by two different kinds of normativity (say, by both prudence and morality, to take a happy example), the same object can have different kinds of value (e.g., moral and aesthetic), and so on. A second, more substantive sort of interaction is *encroachment*. This is the phenomenon wherein reasons or requirements from one domain of normativity make a difference to reasons or requirements in a different domain of normativity.[73] For example, according to *eudaimonic encroachment*, reasons and requirements relating to our own well-being can make a difference to our moral reasons and requirements without themselves being moral.[74] The mechanics of and the debates surrounding encroachment are outside the scope of this article, though they are worth exploring in the future as they relate to political normativity. The point for now is that it is a viable example of how two domains can interact in a fairly strong way while remaining distinct. Of particular interest to us, then, is that the interaction of the moral and political is not necessarily evidence that there is no distinction between the two, nor that political considerations reduce to moral considerations. It is a mistake to automatically infer identity from interaction or to think that cross-domain relevance implies reduction.[75]

Thus, it is consistent with the view on offer that morality plays an important role in political theory and in politics itself. This can happen in a subjective and an objective sense. Regarding the latter, objective moral facts might inform political normativity in the same ways that we just mentioned—in terms of encroachment, for example. Regarding the former, what makes sense to people—and thus what counts as politically normative—is plausibly informed by their moral views. Justifications for power are usually not acceptable to people if they strike them as grossly immoral. But this is true whether or not the moral views in question are correct. In any case, we can remain neutral on the question of whether these kinds of relationships actually obtain; the point

73 Perhaps the most well-known kind of encroachment is *pragmatic encroachment* in epistemology. This is the view that pragmatic facts can make a difference to whether one has reason to form a certain belief. Importantly, they do so without themselves being reasons for belief (i.e., epistemic reasons). For discussion, see Ichikawa and Steup, "The Analysis of Knowledge," sec. 12. See also note 9 above.

74 Lord, "Impartiality, Eudaimonic Encroachment, and the Boundaries of Morality." This is analogous to pragmatic encroachment in epistemology.

75 For a similar point, see also Jubb, "What a Distinctively Political Normativity Is," 365.

is that even if they do, they do not threaten the distinctiveness of the domains.[76] Recall once again the analogy with moral-prudential conflicts: what is conducive to our own well-being is bound up with what we think is moral, given that most of us care about morality. Still, morality is not prudence.

Such interactions can work in the opposite direction as well. That is, political considerations can make a difference to moral considerations without themselves being moral. This opens conceptual space for the possibility that giving due regard to others could involve building political orders that make sense to them as authoritative, without it being the case that MS itself is a moral value. Analogies involving other normative domains are again helpful here. It may be that I owe it to my children to expose them to Miles Davis records or Jane Austen novels, but this does not mean that the value of those works is moral rather than aesthetic. A more plausible view is that their aesthetic value makes a difference to what counts as my giving due regard to my children. Similarly, the fact that MS facts can make a difference to morality does not necessarily mean that such facts are themselves moral. Moralists, then, must do more than point out the existence of these sorts of interactions.[77]

5. The Significance of Political Normativity

To recap, the previous section considered lingering worries that MS is a moral notion. In the end, the upshot is that MS is best understood as a nonmoral notion. It cannot be adequately accommodated within the moral domain. Let us now turn to a different sort of objection. Here, a moralist may grant that there is some nonmoral evaluative standard based around MS but deny that this standard is worth paying attention to. This brings us back to section 1 and to the way we have set up the issue of domains. I define them, following Lord, as systems of evaluations, recommendations, and demands. But as he points out, this means that domains come rather cheaply.[78] Practices like etiquette and grammar and games like chess all have associated systems of evaluations,

76 Some realists seem to dispute this, leading to pessimism about the relevance of the claim of a distinctive political normativity. See, for example, Sleat, "Realism and Political Normativity," 8.

77 In this sense, the kind of political realism associated with the view of political normativity on offer here is what Matt Sleat and Enzo Rossi refer to as the weaker rather than stronger version ("Realism in Normative Political Theory," 690). The stronger version seeks to excise morality from political philosophy altogether; the weaker version allows it to play a role. I follow many realists, such as Edward Hall (*Value, Conflict, and Order*, 13, 185n42), in seeing little value in the stronger version.

78 Lord, "Choosing the Right Companion," sec. 2.1.

recommendations, and demands. It seems clear, then, that not all domains are equally worthy of our attention; certain domains seem especially important. For example, etiquette and chess matter less than morality. The moralist, then, can concede the existence of a political domain of normativity but argue that it is more like etiquette than morality. It is insignificant.

An immediate problem with this view is that it is in tension with the historical and sociological significance of MS that we already noted in section 3. Indeed, it is in tension with numerous considerations that we have raised in building the argument for political normativity. For example, political normativity is necessary for our understanding of certain kinds of political conflicts, and it affects how we think about the value of democracy and public reason. Similar considerations cannot be raised regarding etiquette or the rules of chess. It is also easy to explain why political normativity has this kind of significance: it speaks to a central aspect of human experience. This makes it significant in just the same way that the paradigmatic domains are significant, even if we quibble about their relative significance: they concern what is true, what is beautiful, what gives due regard to others, what makes our lives go well, and so on. These are all matters of great importance, certainly more so than questions about which chess move to make or whether I have set the table correctly. Building political orders that make sense as authoritative to their inhabitants is more than just a game.

There is a more specific way to think about significance, though, that a moralist might have in mind when objecting that the political is insignificant. This is the idea that certain domains are *authoritative*, in that their dictates are inescapable by any rational agent.[79] It is widely held that the epistemic is authoritative in this way, and a longstanding ambition in moral philosophy is to show that the same is true of morality.[80] For many, the notion of authority marks a fundamental distinction—on the one hand, we have domains like the rules of chess, and on the other hand, we have domains like morality that are genuinely action guiding, that *really* tell us what to do.[81] If a defender of political normativity

79 Lord, "Choosing the Right Companion."

80 This is related to the thought that certain domains are overriding, i.e., that whenever there is competition, they win. These notions are not necessarily the same, though. One could hold that both the moral and the epistemic are authoritative without having a view about whether one overrides the other in cases of conflict. On this view, there may simply be genuine practical dilemmas.

81 Baker, "Skepticism About Ought *Simpliciter*," 234; "The Varieties of Normativity," 577–80.

wants to claim significance, they must show that the political is like morality in this way. If it is not, then it is not worth caring about.[82]

A straightforward way to rebut such an objection is to give examples where political reasons ultimately determine what we ought to do. Such examples are bound to be controversial, given that they involve first-order, all-things-considered judgments about what to do when moral and political reasons conflict. Notably, I avoid relying on such judgments; the argument needs to deal only with the underlying structure of the cases at issue. The idea is to show that they involve conflicts that cannot be explained from within a single domain of normativity. The present objection is about whether the political might win out in some such cases.

It seems fair to say that they might, though there is room for disagreement. To illustrate, consider the issue of climate justice. Most environmentalists, philosophers, and activists agree that climate justice places considerable moral demands on well-off societies. For example, such societies should take significant action to curtail greenhouse gas emissions. There are many potential policies that could help comply with such demands: widespread carbon taxes, aggressive measures to phase out combustion engine vehicles, restrictions on high-emission agricultural practices, even limiting air travel, to name just a few examples. Grant for the sake of argument what seems to be the case: these measures do not garner much popular political support in the United States, at least if electoral results are anything to go by. Perhaps there might be certain policies that are slightly more popular, but these alone are less than what is morally required. To this extent, then, the politics that climate justice demands do not make sense as authoritative to a great many people, at least at the time of writing. Would the government be wrong to pursue them anyway at this

82 Some theorists, such as Stephen Finlay, even reserve the term 'normativity' exclusively for phenomena in the latter category, preferring terms like 'norm-relativity' for the former ("Recent Work on Normativity," 331–33). Others distinguish between weak or formal normativity, on the one hand, and authoritative normativity, on the other. I follow Lord ("Choosing the Right Companion," sec. 2.1) and Baker ("The Varieties of Normativity," 568–69) in considering all of this to count as normativity, but I agree with Finlay, Leader Maynard and Worsnip ("Is There a Distinctively Political Normativity?" 778–79), and just about everyone else that some normativity is more important than other normativity. This is what matters in the present context, not the terminology. Those philosophers who endorse this sharp divide here understand the claim of a distinctively political normativity differently from those who are content to call etiquette a domain of normativity. For the former group, a defender of the political needs to show that political demands tell us what to do. For the latter, there simply needs to be some recognizable system of recommendations, evaluations, and demands. This bar is much lower, arguably to the point of triviality. Simply establishing the political in this sense is a pyrrhic victory for realists, especially in the eyes of moralists.

time? Plausibly, yes. But even the fact that this is an open question suggests that morality is not alone in the practical calculus. Distinctly political considerations demand attention as well.

Whether this particular example succeeds, there is a larger point in the background that is crucial to bring to the fore: it is not the case that a defender of a political domain of normativity must uncontroversially show that it wins out in practical deliberation. The demand that we show this rests on an overly simplistic view of authority. There is important space between the poles of domains that "really" tell us what to do and domains that are only formally or weakly normative.

There are two reasons to think this: first, it is not the case that authority entails total dominance over all of our practical and doxastic attitudes.[83] All that it entails is that it always matters in some sense, i.e., that rationality takes it into account. Domains can fully determine some weighted verdict, but they can also matter to determining the weighted verdict without being the only considerations in play.[84] They can exert rational pressure without winning out. Indeed, the examples of ordinary normative conflict given in section 1 exhibit this point. The second reason is that authority can obtain with regard to some practical and doxastic attitudes but not others. For example, morality might fully determine what we ought to intend to do but not what we believe. These two points together are what allow us to properly articulate and account for the complexity of normativity's relationship to our various practical and doxastic attitudes. They allow us to say, for example, that the aesthetic domain fully determines what aesthetic judgment I should make in a particular case, but it lacks this kind of authority for what I ought to intend to do. Aesthetic considerations, though, might factor into the overall weighted verdict regarding what I ought to intend to do.[85] This refined view of authority complicates the simple dichotomy between domains that "really" tell us what to do and those that do not.

This matters for the present discussion because adherence to this dichotomy has allowed claims of a distinctively political normativity to be too quickly dismissed.[86] Leader Maynard and Worsnip, for example, frequently imply that unless political reasons "really" tell us what to do, political normativity is not really normative at all.[87] But this puts things in a misleading way, conflating the

83 On these points, I am indebted to Errol Lord.

84 See Lord's distinction between *verdictive authority* and *contributory authority* ("Choosing the Right Companion," 8–9).

85 Lord, "Choosing the Right Companion," 8. To take another example, morality might fully determine what we ought to intend to do but not what we ought to believe.

86 For discussion, see Hernandez, "Why Isn't There a Distinctively Political Normativity?"

87 See, for example, Leader Maynard and Worsnip, "A Distinctively Political Normativity?" 780, 781–85.

general idea of authority with a more specific kind of it—namely, full authority over the practical. As we have seen, though, authority is a more complex notion. There is more space than moralists let on for a domain like the political to be significant for our practical lives. It can be more significant than etiquette without fully determining the weighted verdicts of the practical.[88]

Giving a full theory of the authority of the political domain is beyond the scope of this article. My goal is only to establish that there is such a domain. The point of raising the issue of authority is to stave off the moralist objection that even if there is a political domain, it is not worth paying attention to. Not only is this objection implausible on its face, but it also rests on the assumption that in order to be worth paying attention to, a domain must have an implausibly strong kind of authority. Insofar as they help themselves to this assumption, moralists stack the deck against political normativity. Even without a full theory of the political's authority in hand, it is important to level the playing field.[89]

6. Conclusion

It is time to take stock. In sections 1 and 2, we motivated the idea of political normativity by arguing that it helps to explain a persistent kind of normative conflict in politics. In section 3, we filled out the idea of political normativity in terms of a central organizing value—namely, MS, the value of a political order making sense to its subjects as authoritative—in such a way as to vindicate the motivation. In section 4, we gave some rejoinders to the moralist objection that we can accommodate MS and its importance wholly from within the moral. And in section 5, we rebutted the worry that MS is only formally or weakly normative.

Regarding the question of political normativity, then, this article's contributions are twofold. On the one hand, I outline a general approach to distinguishing domains of normativity, consisting in the identification of normative conflicts that are best explained by the existence of distinct domains. This approach has been fruitful in other areas of ethical theory and can be applied to the question of political normativity. The second contribution of the article is to pursue this strategy in the case of the political. As I mentioned at the outset, the two contributions are separable. In other words, one can adopt my suggested methodology and come to different conclusions about the substantive

88 We might also find significance outside of the practical, once we consider other attitudes. Perhaps, for example, the political domain fully determines whether it is apt to be alienated from one's political order when it does not make sense.

89 Again, see Hernandez, "Why Isn't There a Distinctively Political Normativity?"

conception of the political domain (though they would have to engage with the arguments for that conception that I give here).[90]

The development of this methodology also serves to push back against a kind of realist worry about the approach I take here. Robert Jubb, in response to Leader Maynard and Worsnip's paper criticizing the idea of a distinctive political normativity, argues that the method of using "abstract general arguments" and "attempt[ing] to demonstrate, in analytically rigorous terms, that moralism makes conceptual mistakes" is unlikely to be fruitful.[91] On his view, the problem has to do with the lack of prospects for agreement on the terms of the argument. There is no shared sense of what counts as a distinctive political normativity, nor what could count as evidence for or against it.[92] This results in theorists talking past each other and failing to make progress. Jubb's suggestion is that we change tack, taking up "a more discursive, allusive method," focused, for example, more on particular contexts.[93] So even though my argument is friendly to the realist project, some realists may chafe against the sort of tools I bring to bear.[94]

I take seriously this concern, but I believe that this article moves forward on precisely the questions that Jubb is pessimistic about. I try to provide a clear idea both of what a distinctively political domain of normativity might look like and how one might argue for and against it. I draw on a relatively uncontroversial notion of normative domains to conceive of the political as a nonmoral domain, analogous to the epistemic, aesthetic, and so on, with a central organizing value. And I work up a theory regarding what counts as evidence for or against it: in short, whether it has explanatory power when it comes to ethical conflict. While this in part involves schematic cases that someone with Jubb's pessimism might object to, it is now open to all of us to develop richer and more realistic cases within the methodology that I propose. One way to think of this is that we need a bit of abstract, conceptual work in certain areas of the realist project, and more contextual and concrete work in others, and that we need those areas to inform each other. I lean toward the former in this article, but this is not incompatible with such an overall picture, and I see no reason

90 In theory, one can also agree with my conception of political normativity without adopting my methodology.

91 Jubb, "On What a Distinctively Political Normativity Is," 363, 361.

92 Jubb, "On What a Distinctively Political Normativity Is," 366. Sleat makes a similar point as well ("Realism and Political Normativity," 8).

93 Jubb, "On What a Distinctively Political Normativity Is," 367.

94 I thank a referee at this journal for suggesting that I reflect on Jubb's concern here. It is also worth keeping in mind that realism is a big tent. See Sleat, "Realism and Political Normativity," sec. 3.

to think that that picture is incoherent or unattractive. Indeed, it seems to me quite in the spirit of Williams's work, which fits neatly into neither a purely analytic and abstract approach nor a purely historicized and concrete approach, and has subsequently been productively engaged with by theorists on either side. My point here, then, is simply that such conversation can be dialectically constructive, and it is not clear to me that the appropriate move at this juncture is for both sides to go their separate ways. The optimistic conclusion is that more engagement will continue to be fruitful.

With that being said, we can now conclude with a brief summary. In this article, I have argued that there is a distinct, nonmoral, political value in the achievement of political power that makes sense to its subjects as authoritative. This value is at the center of the political domain of normativity. Independently of that claim, I have also tried to contribute to the political normativity debate by clarifying what it is about. It is about, I have suggested, whether a distinctly political domain is necessary to make sense of the normative richness of politics. I have shown how political normativity understood in terms of MS fits this bill: it explains a persistent kind of conflict in political theory and practice, between what morality demands and what makes sense to people as an authoritative order of power. This has potential implications for a number of debates in political philosophy, some of which I explored in section 3. While I leave open the question of exactly how the demands of politics weigh up against morality, the point for now is just that they are distinct demands. When it comes to politics, we have reasons not only to build the most just society but to build a society that is ours, that makes sense to us, given who we take ourselves to be and what problems we take ourselves to have. If the central claim here is correct, then there is a distinct kind of value in this creative project.[95]

State University of New York at Oswego
michael.gadomski@oswego.edu

95 I am grateful to audiences at the American Philosophical Association (APA) Central Division Meeting, the Rocky Mountain Ethics Congress, and the Penn Normative Philosophy Group, and to two anonymous referees and an editor at this journal. Thanks especially to Errol Lord for many conversations about this topic and to Amanda Greene and Seungsoo Lee for extensive written comments. I also thank Samuel Freeman, Kok-Chor Tan, Rob Hughes, John Roman, Paul Musso, Brian Berkey, Steph Wesson, Max Lewis, and Shawn Hernandez.

REFERENCES

Annas, Julia. "Prudence and Morality in Ancient and Modern Ethics." *Ethics* 105, no. 2 (1995): 241–57.

Arneson, Richard J. "Democracy Is Not Intrinsically Just." In *Justice and Democracy: Essays for Brian Berry*, edited by Keith Dowding, Robert E. Goodin, and Carole Pateman. Cambridge University Press, 2004.

Bagg, Samuel. *The Dispersion of Power: A Critical Realist Theory of Democracy*. Oxford University Press, 2023.

———. "The Power of the Multitude: Answering Epistemic Challenges to Democracy." *American Political Science Review* 112, no. 4 (2018): 891–904.

———. "Realism Against Legitimacy: For a Radical, Action-Oriented Political Realism." *Social Theory and Practice* 48, no. 1 (2022): 29–60.

Baker, Derek. "Skepticism About Ought *Simpliciter*." In *Oxford Studies in Metaethics*, vol. 13, edited by Russ Shafer-Landau. Oxford University Press, 2018.

———. "The Varieties of Normativity." In *The Routledge Handbook of Metaethics*, edited by Tristram McPherson and David Plunkett. Routledge, 2017.

Burelli, Carlo. "Political Normativity and the Functional Autonomy of Politics." *European Journal of Political Theory* 21, no. 4 (2022): 627–49.

Burelli, Carlo, and Chiara Destri. "The Sources of Political Normativity: The Case for Instrumental and Epistemic Normativity in Political Realism." *Ethical Theory and Moral Practice* 25, no. 3 (2022): 397–413.

Christiano, Tom, and Sameer Bajaj. "Democracy." *Stanford Encyclopedia of Philosophy* (Summer 2024). https://plato.stanford.edu/archives/sum2024/entries/democracy/.

Copp, David. "The Ring of Gyges: Overridingness and the Unity of Reason." *Social Philosophy and Policy* 14, no. 1 (1997): 86–101.

Driver, Julia, "The Conflation of Moral and Epistemic Virtue." In *Moral and Epistemic Virtues*, edited by Michael Brady and Duncan Pritchard. Blackwell, 2003.

Eaton, Marcia Muelder. "Integrating the Aesthetic and the Moral." *Philosophical Studies* 67, no. 3 (1992): 219–40.

Enoch, David. "Against Public Reason." In *Oxford Studies in Political Philosophy*, vol. 1, edited by David Sobel, Peter Vallentyne, and Steven Wall. Oxford University Press, 2015.

Erman, Eva, and Nicklas Möller. "The Role of Moral Norms in Political Theory." *Topoi* 44, no. 1 (2025): 27–38.

Finlay, Stephen. "Recent Work on Normativity." *Analysis* 70, no. 2 (2010): 331–33.

Fletcher, Guy. *Dear Prudence: The Nature and Normativity of Prudential Dis-*

course. Oxford University Press, 2021.
Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007.
Geuss, Raymond. *Philosophy and Real Politics.* Princeton University Press, 2008.
Greene, Amanda R. "Is Political Legitimacy Worth Promoting?" *Nomos* 61 (2019): 65–101.
———. "Legitimacy Without Liberalism: A Defense of Max Weber's Standard of Political Legitimacy." *Analyse und Kritik* 39, no. 2 (2017): 295–323.
Hall, Edward. *Value, Conflict, and Order: Berlin, Hampshire, Williams, and the Realist Revival in Political Theory*. University of Chicago Press, 2020.
Hannon, Michael, and Jeroen de Ridder, eds. *The Routledge Handbook of Political Epistemology*. Routledge, 2021.
Hernandez, Shawn. "Why Isn't There a Distinctively Political Normativity?" Unpublished manuscript.
Ichikawa, Jonathan Jenkins, and Matthias Steup. "The Analysis of Knowledge." *Stanford Encyclopedia of Philosophy* (Fall 2024). https://plato.stanford.edu/archives/fall2024/entries/knowledge-analysis/.
Jubb, Robert. "On What a Distinctively Political Normativity Is." *Political Studies Review* 17, no. 4 (2019): 360–69.
Leader Maynard, Jonathan. "What Is Political Moralism?" *Topoi* 44, no. 1 (2025): 13–26.
Leader Maynard, Jonathan, and Alex Worsnip. "Is There a Distinctively Political Normativity?" *Ethics* 128, no. 4 (2018): 756–87.
Lord, Errol. "Choosing the Right Companion: On the Authority of Epistemic, Aesthetic, and Moral Normativity." In *The Future of Normativity*, edited by Simon Kirchin. Oxford University Press, 2025.
———. "Impartiality, Eudaimonic Encroachment, and the Boundaries of Morality." In *Oxford Studies in Normative Ethics*, vol. 11, edited by Mark Timmons. Oxford University Press, 2021.
Mill, John Stuart. *Utilitarianism*. 2nd ed. Edited by George Sher. Hackett Publishing, 2001.
Miller, David. *Is Self-Determination a Dangerous Illusion?* Polity Press, 2019.
Muñoz, Daniel. "Three Paradoxes of Supererogation." *Nous* 55, no. 3 (2021): 699–716.
Musso, Paul. *The Historical and Contemporary Significance of Anscombe and Foot's Metaethical Thought*. PhD diss., University of Pennsylvania, 2022.
Nelson, Maggie. "Art Song." In *On Freedom: Four Songs of Care and Constraint*. Graywolf Press, 2021.
Piketty, Thomas. *Capital and Ideology*. Translated by Arthur Goldhammer. Harvard University Press, 2020.

Quong, Jonathan. "Public Reason." *Stanford Encyclopedia of Philosophy* (Summer 2022). https://plato.stanford.edu/entries/public-reason/.

Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13, no. 2 (1984): 134–71.

Rawls, John. *The Law of Peoples: With "The Idea of Public Reason Revisited."* Harvard University Press, 1999.

———. *Political Liberalism*. Columbia University Press, 1996.

Rodríguez-Alcázar, Javier. "Beyond Realism and Moralism: A Defense of Political Minimalism." *Metaphilosophy* 48, no. 5 (2017): 727–44.

Ross, W. D. *The Right and the Good*. Edited by Philip Stratton-Lake. Oxford University Press, 2002.

Rossi, Enzo. "What Can Epistemic Normativity Tell Us About Politics? Ideology, Power, and the Epistemology of Radical Realism." *Topoi* 44, no. 1 (2024): 77–88.

Rossi, Enzo, and Matt Sleat. "Realism in Normative Political Theory." *Philosophy Compass* 9, no. 10 (2014): 689–701.

Sagar, Rahul, and Andrew Sabl, eds. *Realism in Political Theory*. Routledge, 2018.

Sleat, Matt. "Politics Recovered: On the Revival of Realism in Contemporary Political Theory." In *Politics Recovered: Realist Thought in Theory and Practice*, edited by Matt Sleat. Columbia University Press, 2018.

———, ed. *Politics Recovered: Realist Thought in Theory and Practice*. Columbia University Press, 2018.

———. "Realism and Political Normativity." *Ethical Theory and Moral Practice* 25, no. 3 (2021): 465–78.

———. "What Is a Political Value? Political Philosophy and Fidelity to Reality." *Social Philosophy and Policy* 33, nos. 1–2 (2016): 252–72.

Sommavilla, Vitor. "A Constructivist Account of Political Normativity." *Topoi* 44, no. 1 (2024): 65–75.

Tiffany, Evan. "Deflationary Normative Pluralism." *Canadian Journal of Philosophy* 37, no. 5 (2007): 231–62.

Valentini, Laura. "Justice, Disagreement, and Democracy." *British Journal of Political Science* 43, no. 1 (2012): 177–99.

Walzer, Michael. "The Moral Standing of States: A Response to Four Critics." *Philosophy and Public Affairs* 9, no. 3 (1980): 209–29.

Wendt, Fabian. "Compromise and the Value of Widely Accepted Laws." In *Compromise and Disagreement in Political Theory*, edited by Christian Rostbøll and Theresea Scavenius. Routledge, 2017.

Williams, Bernard. *In the Beginning Was the Deed: Realism and Moralism in Political Argument*, edited by Geoffrey Hawthorn. Princeton University Press, 2005.

———. "Philosophy as a Humanistic Discipline." In *Philosophy as a Humanistic Discipline*, edited by A. W. Moore. Princeton University Press, 2006.
———. "Realism and Moralism in Political Theory." In *In the Beginning Was the Deed: Realism and Moralism in Political Argument*, edited by Geoffrey Hawthorn. Princeton University Press, 2005.
Wolf, Susan. "Moral Saints." *Journal of Philosophy* 79, no. 8 (1982): 419–39.
Worsnip, Alex. "Eliminating Prudential Reasons." In *Oxford Studies in Normative Ethics*, vol. 8, edited by Mark C. Timmons. Oxford University Press, 2018.

Journal of Ethics and Social Philosophy
vol. 30, no. 6 · october 2025
https://doi.org/10.26556/jesp.v30i6.4418

CAN WE EAT ANIMALS WHOSE EXISTENCE DEPENDS ON IT?

Karri Heikkinen

One common defense of eating meat appeals to the idea that if we did not farm animals for human consumption, these animals would never exist. Consequently, at least if the animals have sufficiently good lives, those eating meat may in fact be doing them a service. A classic statement of this argument comes from D. G. Ritchie, who writes:

> If all the world were Jews, it has been well said, there would be no pigs in existence; and if all the world were Vegetarians, would there be any sheep or cattle, well cared for, and guarded against starvation?[1]

Leslie Stephen echoes the sentiment, writing:

> The pig has a stronger interest than anyone in the demand for bacon. If all the world were Jewish, there would be no pigs at all. He has to pay for his privileges by an early death; but he makes a good bargain of it.[2]

This point has since become so well known that anyone interested in the ethics of eating meat is likely to have heard it in some form. But should vegetarians be moved by it?

As far as I am aware, most vegetarians have not been moved. This, I believe, is largely due to the troubling ways the very same reasoning can be applied to humans. A contemporary opponent of Ritchie and Stephen, the animal advocate Henry S. Salt comments on the argument as follows:

> It has, in fact, been the plea of the slave-breeder; and it is logically just as good an excuse for slave-holding as for flesh-eating. It would justify parents in almost any treatment of their children, who owe them, for the great boon of life, a debt of gratitude which no subsequent services can repay. We could hardly deny the same merit to cannibals, if they were to breed their human victims for the table.[3]

1 Ritchie, *Natural Rights*, 110.

2 Stephen, *Social Rights and Duties*, 236.

3 Salt, *The Humanities of Diet*, 221.

Something very close to Salt's reply, I think, is also familiar to those working on the ethics of eating meat. Importantly, it seems to make our original argument—or the "logic of the larder," as Salt memorably calls it—obsolete. For if the argument is to work, it must be complemented with some argument justifying the differential treatment of humans and animals. But if that can be done, then it looks unlikely that we still need our original argument. Instead, we could presumably just appeal directly to whatever justifies this differential treatment and thereby justify omnivorism. Unsurprisingly then, much of the contemporary debate on the permissibility of eating meat revolves around the more fundamental issue of the moral status of animals.

In the last couple of decades, however, the picture has grown more complicated. Many consequentialists, despite the influence that consequentialism has on the contemporary animal advocacy movement, now feel the force of the argument. For example, Peter Singer, the author of *Animal Liberation*, now concedes that a limited form of omnivorism may be permissible if it leads to the existence of happy animals, and any animals who die are replaced with new happy animals.[4] This is because, according to Singer's utilitarian view, the permissibility of omnivorism ultimately depends on whether it produces the impersonally best outcomes.[5] If omnivorism leads to the existence of sufficiently many sufficiently happy animals, their happiness may ultimately make the world impersonally better.[6] Something very much like this view is also suggested by R. M. Hare.[7]

Furthermore, even many philosophers with a consequentialist orientation who reject omnivorism still take the argument to be worth struggling with. For example, Tyler John and Jeff Sebo argue that eating meat generally leads us to form attitudes that undervalue animals; therefore, consequentialists still have indirect reasons to resist omnivorism.[8] In a similar vein, Gaverick Matheny and Kai Chan claim that even if we accept the basic idea behind the argument, buying meat is very unlikely to be the most effective way to transform our resources into welfare.[9] Tatjana Višak, on the other hand, argues that utilitarians need not accept the "logic of the larder" since they need not accept the idea

4 Singer, *Animal Liberation*.

5 Singer, *Practical Ethics*, ch. 5.

6 The utilitarian interpretation of argument suggested by Ritchie and Stephen is sometimes called the *replaceability argument*. For a comprehensive overview of the history of this version of the argument, see Delon, "Replaceability Argument in the Ethics of Animal Husbandry."

7 Hare, "Why I Am Only a Demi-Vegetarian."

8 John and Sebo, "Consequentialism and Nonhuman Animals."

9 Matheny and Chan, "Human Diets and Animal Welfare."

that being brought into existence is a benefit.[10] To give one more example, Jeff McMahan, despite having considered and rejected the argument, later admits in an interview that he still does not know if omnivorism is wrong.[11]

Of course, some readers may take consequentialism's failure to decisively vindicate vegetarianism as a reason to reject consequentialism rather than vegetarianism. But recently, this position has also come under threat. This is due to Abelard Podgorski, who takes up the kind of argument suggested by Ritchie and Stephen and develops it into what he calls the *diner's defense*.[12]

What makes Podgorski's version of the argument particularly impressive is that he makes very significant concessions to vegetarians at the outset: he grants animals the exact same moral status as humans and accepts that animals are harmed when they are killed. Furthermore, he does not rely on any form of impersonal consequentialism, meaning that, unlike Singer, Podgorski does not think that the harm done by killing animals can somehow be offset by replacing them with new happy animals. Yet Podgorski manages to argue that eating certain kinds of meat can be permissible essentially because by buying meat, we bring into existence animals who benefit from their existence overall. This development means that contrary to what vegetarians following Salt might be inclined to say, the "logic of the larder" cannot be easily dismissed by pointing to its implications for humans nor by labelling it as an obscure implication of consequentialist theorizing.

In this article, I pick up the debate on behalf of vegetarians. The article has two aims. In sections 1 and 2, I aim to work out the best version of the argument we started with. While I make significant use of Podgorski's recent development of the idea, I also argue that he misses some of its most striking implications. One of these implications deserves to be mentioned straight away: if we take the argument seriously, then consuming dairy seems to be significantly morally worse than eating meat. I then ask whether we should accept the revised argument, and I argue that we should not. This is for three reasons, all of which I explain in depth in the relevant sections. In section 3, I argue that the resulting defense of omnivorism makes the so-called nonidentity problem unavoidable. In section 4, I show how a significant part of its intuitive force seems to problematically rely on contingent features of our world. Finally, in section 5, I argue that even the best version of the argument still seems to

10 Višak, *Killing Happy Animals*.

11 McMahan, "Eating Animals the Nice Way." For the interview, see Goldhill, "An Oxford Philosopher's Moral Crisis Can Help Us Learn to Question Our Instincts."

12 Podgorski, "The Diner's Defence."

have problematic implications when generalized to humans, even beyond the nonidentity problem.

1. The Diner's Defense

I credit Podgorski for significantly developing the kind of argument that Ritchie and Stephen suggest, so a natural place to start is to look at Podgorski's view in more detail. Podgorski aims to provide a qualified defense of consuming the meat of farmed animals (and, by extension, as we shall see in section 2, other products made using animals). He puts in place three qualifications. First, he is defending only the consumption of meat, not its production (so, for example, he is not defending farming). As we shall soon see, this restriction is essential for avoiding some of the criticism that Salt leverages against Ritchie and Stephen. Second, the animals in question must have lives that are worth living overall, meaning that the argument may not apply to all or even most of the meat currently on the market. Third, Podgorski's argument is only that there is no *harm-based* reason against consuming meat, meaning that we might still have other reasons to refrain from this practice (for example, the negative environmental effects of beef production). For readability, I do not repeat these qualifications. Unless explicitly stated otherwise, readers should assume that whenever I talk about, for example, Podgorski's view permitting the consumption of animal meat, I mean to say that the view has this implication when considering only harm-based reasons and only farmed animals with lives worth living.

Two things make Podgorski's argument particularly worthy of attention. First, the most influential arguments against moral omnivorism in the literature invoke harm-based reasons in one way or another, for example by claiming that eating animals violates their rights or causes them suffering in exchange for mere gustatory pleasure for us.[13] So even though Podgorski's argument is limited, it nevertheless challenges very influential views. Second, as already mentioned, Podgorski makes multiple exceptionally significant concessions to vegetarians at the outset. Most importantly for the purposes of this article, Podgorski grants to his opponents that farm animals have the same moral status as humans and that death is bad for these animals.

With these preliminaries in place, we can move on to the argument itself. Podgorski begins by considering two ideas that often come up in discussions of vegetarianism, which he calls *existence-dependence* and *indirectness*. Roughly, the

13 For some of the most influential arguments in this vein, see Singer, *Animal Liberation*; Regan, *The Case for Animal Rights*; and Norcross, "Puppies, Pigs, and People."

former idea is what we started with—namely, that the animals we eat would not have existed at all were it not for us farming them with the intention of eating them. The latter idea concerns the fact that consumers do not typically kill the animals they eat but rather leave this to the producers of meat. While neither of these ideas is sufficient to justify omnivorism on its own, Podgorski thinks we can combine them into an argument that is up to the task.

According to Podgorski, there is an important kernel of truth in the existence-dependence idea. He summarizes the lesson as follows:

> An animal or person does not have a harm-based complaint against an act that causes them to exist, if no alternative act would have led to their existence, and their existence is good for them.[14]

This holds even if the person's or animal's life involves severe harms, as long as their life is overall positive. Note, however, that this principle applies only to acts that are necessary for the person's or animal's existence: after they are born, they can have a complaint against any subsequent act, distinct from the act that brought them into existence, that makes their life worse than it could have otherwise been.

It can be objected that existence itself cannot be said to be good for someone or to benefit them. In fact, Salt already thought of this objection over a century ago. He writes:

> The fallacy lies in the confusion of thought which attempts to compare existence with non-existence. A person who is already in existence may feel that he would rather have lived than not, but he must first have the *terra firma* of existence to argue from; the moment he begins to argue as if from the abyss of the non-existent, he talks nonsense.[15]

In response to this objection, it might be pointed out that it seems eminently possible to imagine a life so bad that it would be worse than nonexistence; and if that seems possible, then it is hard to say exactly why a life cannot be better than nonexistence as well. Furthermore, many philosophers now accept the idea that an existence can be good or bad in a noncomparative way, thus avoiding the problem of comparing existence with nonexistence.[16]

The best response, however, is to simply avoid this problem by modifying Podgorski's argument slightly. Instead of speaking about existence being good for an animal or person, we can just say that an animal or a person does not

14 Podgorski, "The Diner's Defence," 71.

15 Salt, *The Humanities of Diet*, 221.

16 See, for example, McMahan, "Causing People to Exist and Saving People's Lives."

have a complaint against an action that causes them to exist if that existence is *not worse* than nonexistence. This weaker principle is all that is needed for the argument to go through, and it is very difficult to deny that a life can be not worse than nonexistence: if that was not true, then we would hardly ever be justified in creating anyone at all.[17]

This principle becomes crucial when it is combined with what Podgorski takes to be the important takeaway from the idea of indirectness. He states this second lesson as follows:

> Producers of animal meat have control over whether animals suffer and die that comes apart from their control over whether those animals exist. Consumers of animal meat do not.[18]

The idea here is that the indirect relationship between consumers and the suffering of animals grounds an asymmetry between the consumers and producers of animal products. Producers of animals could in principle decide to breed animals and then refrain from putting them into small cages or killing them so that their meat can be sold. But when a consumer buys meat in a shop, they buy meat from an animal that has obviously already been slaughtered. Therefore, it is not the case that when a consumer buys meat, they harm existing animals or that if they do not buy meat, existing animals are spared from harm. Instead, the purchase sends a signal to the producers of meat that there is demand, and this leads the producers to breed new, additional animals. From the point of view of the consumer, the only way to stop the suffering and untimely death of these potential future animals is by not buying meat, but this would mean that these very animals never come to exist in the first place. So we get the asymmetry: consumers of animal meat can decide only whether future farm animals exist or not, whereas producers of animal meat can decide both whether these animals exist and, independently, whether these animals suffer and die on their farms.

17 Višak claims that bringing an animal into existence can neither harm nor benefit it (*Killing Happy Animals*). I find this view very difficult to defend given the strong intuition that it is possible for a life to be so bad that it would be better to not exist than to experience that kind of life. Furthermore, while I do not have the space to engage with her view here, I believe that Višak's attempt to make this claim compatible with our intuitions in cases that instantiate the so-called nonidentity problem (see Parfit, *Reasons and Persons*, ch. 16) does not succeed. However, even if we leave these points aside, the response I give in the main text should suffice. It is surely the case that if bringing an animal into existence can neither harm nor benefit an animal, then an animal is not made worse-off when it is brought into existence. This is all we need for the argument at hand to get off the ground.

18 Podgorski, "The Diner's Defence," 71.

Putting these two claims together, we get what Podgorski calls the diner's defense. Because of indirectness, a consumer of animal products faces a choice with only the following two options:

1. Do not eat meat, with the consequence that some possible future animals never exist at all.
2. Eat meat, with the consequence that these same animals do exist and live lives (marginally) worth living.

Because of existence-dependence, the animals have no complaint against option 2, for this choice is necessary for their existence, and this existence is not worse than nonexistence. Therefore, the argument goes, it is permissible for the consumer to choose option 2. After all, choosing this option does not make anyone worse-off than they would have otherwise been, for choosing otherwise would mean the relevant animals never exist at all.

Podgorski holds that this argument can be used to defend only diners: it cannot be used to defend producers of animal meat. This is because a producer, unlike a consumer, faces something like the following options:

1. Do not breed the animals, meaning that the animals never exist at all.
2. Breed the animals, put them into small cages, kill them as soon as they are ready for eating, and so on, meaning that the animals live lives marginally worth living.
3. Breed the animals but then let them roam free and refrain from slaughtering them.

Here, option 2 is not permissible because the farmer could instead choose option 3 and give the same animals a better life. So the animals have a complaint against the farmer who puts them into small cages and kills them, because the farmer could stop these things from happening without stopping the animals from existing. But the animals have no complaint against diners, as diners can choose only whether the animals come to exist in the first place.

It is important to note how significant a role the asymmetry between producers and consumers of animal meat plays in the argument. Recall how Salt, in the passage quoted at the beginning of this article, complains that the kind of reasoning we are discussing here could just as well be used by human parents to justify any kind of treatment towards their children. But Podgorski in fact explicitly considers the generalization of his arguments to humans, using an example that responds directly to this worry.[19] I paraphrase the example as follows. Imagine that you are considering whether to purchase a chocolate bar.

19 Podgorski, "The Diner's Defence," 72.

A perfectly reliable oracle tells you that if you purchase the chocolate, it will trigger a complicated chain of events that results in a boy named Theodore being born. Theodore will be beaten and treated cruelly by his family, and he will die young. Nevertheless, Theodore will have a life worth living, and he will be glad to be alive. If you do not purchase the chocolate, Theodore will never be born.

In this example, Theodore is meant to be analogous to the animals that are born as a consequence of us purchasing meat, and his abusive family is meant to be analogous to producers of meat. And yet, Podgorski holds, it is intuitively permissible to buy the chocolate in this example.

What this example shows is that Salt's objection, at least when taken at face value, is too quick: as Podgorski makes clear, both people and animals can have complaints against any action that makes them worse-off than they would have otherwise been *after* they are born. When it comes to eating animals, this means that the animals do in fact have complaints against many things that the producers do, including slaughter. It is just that consumers do not have the power to change these things, without also changing which animals come to exist, meaning that the animals do not have a complaint against the consumers.

The key takeaway from this exchange is the following. We want to find a way to use the existence-dependence argument to justify eating meat without generating implausible results with respect to humans. What we see here is that the asymmetry between producers and consumers achieves exactly this: it allows us to say how putting animals into small cages and killing them is wrong, just as it would be with humans, even though purchasing meat is not.

In sum then, the best version of the argument we started with must in fact make two key claims: first, animals do not have a complaint against actions that cause them to live lives not worse than nonexistence; and second, there is a causal asymmetry between those who buy and eat meat and those who produce it. Now, readers familiar with population ethics might notice a link between this view and the so-called nonidentity problem. I discuss this issue in detail in section 3. Before that, however, we must clarify an important aspect of the argument as it stands: Is it really the case that consumers of animal products have no influence on what happens to already existing animals?

2. The Scope of the Argument

We have seen that for the argument that we started with to matter, it must be compatible with something like the assumption of equal moral status for humans and animals, and for this to be plausible, we must differentiate between the creation of new animals and the harming of existing ones. But it seems to me that even if we accept this distinction, there are many animal products that still cannot be permissibly consumed, because we do in fact harm existing animals when we buy them. A good example of how this might happen is cow's milk, or dairy.

One sometimes underappreciated fact about producing dairy is that cows do not automatically produce milk after a certain age. Rather, milk production involves forcibly impregnating a female cow. Dairy cows need to reach an age of about two years before giving birth for the first time, after which they can typically produce milk until they reach an age between four and six years.[20] This means that, unlike with meat, a consumer buying dairy sends a signal that usually has time to reach farmers during the productive lifespan of *already existing* cows. Consequently, producers are likely to respond to the continued demand for milk by impregnating already existing cows multiple times rather than by continuously bringing new dairy cows into existence.[21]

What all this means is that even if we fully accept the defense of omnivorism that we have discussed so far, buying dairy is likely to be impermissible because doing so harms existing animals. Impregnating a cow after it has already been alive for a couple of years is obviously not necessary for that cow's existence. Furthermore, given that we are talking about already existing animals, a farmer acting out a consumer's will does not mean that the consumer is off the hook: just as paying someone to kill your enemy on your behalf does not absolve you of responsibility, so too paying someone to impregnate a cow on your behalf does not mean that you get to drink milk with a clear conscience.

It can be objected that forcibly impregnating a cow is not harmful to the cow in the same way it is for a human. However, this objection seems to give up the explicit assumption granted to vegetarians that animals have the same moral status as humans. Even if we accept, as I think we should, that humans face much graver psychological harm from such an act of violence than animals do, it is very difficult to see how we could uphold the assumption of equal moral status yet maintain that forcibly impregnating a cow is fully permissible. Furthermore, keep in mind that if we start relaxing this key assumption, then

20 De Vries and Marcondes, "Review."

21 To keep things simple, I do not discuss the so-called inefficacy problem. The inefficacy problem, also known as the problem of collective harm or collective impact, arises when many people acting in some way collectively causes a great harm (or benefit), but no individual action makes a difference. It seems these actions must be wrong (or *pro tanto* required), but given they do not make a difference, it is not easy to explain why. However, I think this problem can be overcome. For an overview of the problem and some potential solutions, see Nefsky, "Collective Harm and the Inefficacy Problem." For further discussion of potential solutions, see Kagan, "Do I Make a Difference?"; Nefsky, "How You Can Help, Without Making a Difference"; and Hedden, "Consequentialism and Collective Action."

there are more direct ways to justify omnivorism than the argument we are investigating here.

Another potential objection is to deny the empirical claim that milk consumption causes harm to existing animals. A natural way to do this is to claim that if we did not buy milk, farmers would simply send dairy cows to be slaughtered for meat. Insofar as this seems even worse than continuing life as dairy cows, it might be thought that we do the cows a favor by keeping them alive with our dairy purchases. In essence, the idea goes, we are paying farmers to keep producing dairy so that the cows at least avoid slaughter.

This reply, however, is ineffective in at least two ways. First, note that in the case where a dairy cow gets slaughtered for meat, there now seems to be a significant risk that this cow can have a complaint against those who buy meat. This is because those buying meat have not caused the existence of the dairy cow that gets sent to slaughter, but rather, they have affected the fate of an animal that exists independently of those purchasing choices. This is unlike in the normal case of meat consumption, where the demand for meat does not make it the case that, as Podgorski puts it, "animals destined for better things are redirected onto someone's plate."[22] However, in the case of our redundant dairy cow, there being demand for meat does mean that our consumption choices lead to an existing animal, who does not owe its existence to meat consumption and who might have otherwise been able to continue its life, to be sent to the slaughterhouse.

Second, even if we accept the idea that we are, in a rather surprising way, saving dairy cows from an even worse fate by buying milk, this seems to imply that we are now committed to a duty to save farmed animals from at least some kinds of treatment. And if that is the case, then it is an oddly convenient coincidence if these duties are such that they require buying dairy but not, for example, nonviolently blocking animals from being transported to slaughterhouses. Furthermore, if we accept the idea that we may sometimes be required to intervene, then it seems clear even in the case of meat that while an animal may have no complaint against the act of buying meat that led to its existence, it can have a complaint against people who *allow* the producer to kill it. Indeed, if we have a duty to save other humans from dying, and animals and humans have equal moral status, then it is not easy to see what justifies our inaction in a situation where billions of animals are killed for meat every year. Even though these points are not inconsistent with the argument for omnivorism that we have been discussing, I suspect they reduce the appeal that this argument has for those who wish to eat meat.

22 Podgorski, "The Diner's Defence," 71.

Where does all this leave us? The specifics depend on several empirical questions regarding what happens to animals as a consequence of various consumer decisions and other possible interventions, as well as the more detailed view we adopt about our duties of rescue. While I do not have the space here to engage with these questions further, two things seem clear to me. First, it seems clear that even the most developed version of the argument we started with cannot justify the consumption of all animal products, even if these animals have good lives. This includes dairy, as well as any other product such that its consumption leads to already existing animals being subjected to treatment that we find unacceptable for humans. This result is interesting and potentially practically very important because many people consider the next best option after veganism, in the sense of refraining from the consumption of all animal products, to be vegetarianism, in the sense of refraining from eating meat while eating products that include dairy, such as cheese. However, if something like the argument Podgorski and his predecessors propose is correct, then vegetarians have it backwards: the thing to forgo is not meat but rather dairy.

That said, it is important to note that some types of animal-meat products are also such that when a consumer buys them, harm is inflicted on already existing animals. This happens in the same way milk production harms already existing cows: we artificially inseminate animals and force them to give birth. Veal, a side product of the dairy industry, is a clear example, but we should also note that in many places around the world, almost all pigs are bred artificially.[23] Breeding pigs have a lifespan of about one and a half to two years, meaning that consumer demand should have enough time to translate into existing pigs having a higher number of litters during their lifetimes. However, practices differ significantly across the world and across different animals. This means that those wishing to justify their consumption of animal products by appealing to something like the argument we have been discussing must not only make sure the animals they consume have lives worth living but also carefully research the way in which these animals come to exist.

Second, it should be noted that unlike many other arguments for omnivorism, such as those that deny the assumption of equal moral status for humans and animals, the argument at hand often implies that if we can, we should stop animals from being killed. While this result may seem obvious once we see how it comes about, it is also so surprising that it is worth spelling out. For example, it could be that there comes a day in the future when we might wake up and permissibly buy a sausage roll for breakfast while also being

23 Zuidema et al., "An Exploration of Current and Perspective Semen Analysis and Sperm Selection for Livestock Artificial Insemination."

obliged to vote in favor of banning the slaughtering of animals for meat in a referendum taking place later that day. And if a reader is inclined to believe in more stringent duties of rescue, then voting may not be enough: after filling ourselves up with our sausage rolls, we may be required to storm a slaughterhouse to liberate the animals there. This result, I wager, is not what most defenders of omnivorism wish to achieve.[24]

3. The Nonidentity Problem

Let us take stock of what we have seen so far. As I pointed out in the introduction, the argument suggested by Ritchie and Stephen must be made compatible with the assumption of equal moral status for humans and animals for the argument to make a difference in the debate on omnivorism. As we learn from Podgorski, the way to make this possible is to draw a sharp distinction between acts that determine whether an animal comes to exist and acts that make existing animals better- or worse-off. But from that, I argue, two important limitations follow: first, the resulting defense of omnivorism cannot justify the consumption of all animal products even when the animals have lives worth living; and second, despite being permitted to buy meat, even consumers are often obliged to take action to stop existing animals from being killed.

Having worked out the argument, we can now ask: Should we accept it? I believe we should not, and in the rest of this article, I mount three objections to that effect. In this section, I explain how the picture of morality that underlies our defense of omnivorism has counterintuitive implications in a wide range of cases that instantiate the so-called nonidentity problem. In the section after that, I argue that the argument also has counterintuitive implications when we imagine away the empirical asymmetry between producers and consumers that we had to posit to make its generalization to humans plausible, even if we focus just on animals for the moment. Finally, I also show that despite all the work done to improve it, the argument still has counterintuitive implications when applied to humans.

To see how the argument at hand leads us to the nonidentity problem, we must note that this argument relies on what we can call a *person-affecting* picture

24 Outside of agriculture, there are many other practices where the consumer may cause harm to existing animals, such as hunting, bullfighting, keeping animals in zoos, and using animals in medical research. If we take Podgorski's view seriously, then many of these practices may be impermissible, even when the animals have lives worth living. However, these cases also raise further empirical and philosophical complications, such as the question of whether living in captivity might in fact spare an animal from even more suffering in the wild. Thus, I leave the investigation of these practices for another time.

of morality. According to this picture, there are no victimless wrongs: for an action to be wrong, it must always wrong some particular individual—that is, the action must make a person worse-off than they would have otherwise been or in some other way give them a complaint against the agent. The gist of the diner's defense is to show that when it comes to consumers buying and eating animal meat, this condition is not met: a consumer does not make any animal worse-off than it would have been; they merely cause a new animal with a life worth living to be born.

The person-affecting picture has some very attractive features. One such feature particularly relevant to our topic is that it can offer a very natural explanation for the so-called *procreation asymmetry*. Roughly speaking, this is the idea that even though we have strong moral reasons to avoid creating beings with very bad lives, there is no moral reason to create beings with very good lives. Under the person-affecting picture, if we create someone with a life so bad that it would be better for them to have never existed, then our action wrongs that person. On the other hand, if we *do not* create someone who would have had a very good life, then there is no one who is wronged by our choice because the person that could have had a very good life simply never exists.

This attractive implication may make the person-affecting picture look like a no-brainer. However, views like this also have some serious problems. For our purposes here, the key issue is the so-called nonidentity problem. This problem, made famous by Derek Parfit's discussion of future generations, arises from the fact that many of our actions influence not only the welfare and number of beings who come to exist but also their identity.[25] Perhaps the easiest way to grasp the problem is through an example.

Imagine that we need to decide between two climate policies, *A* and *B*. Policy *A* involves a massive decarbonization of the economy, across all sectors of society. Policy *B* makes some moderate efforts to reduce climate change but mostly continues business as usual until we run out of fossil fuels. Let us assume both that policy *B* is better for those currently alive because it means the current generation avoids the costs of the transition and that policy *B* will not make things in the future so bad that it would be better for future generations to not exist at all. However, let us also assume that policy *A* would make things go enormously better for every subsequent generation who comes to exist after the next fifty years.[26]

It seems intuitive that, morally, we ought to choose policy *A*: after all, there is so much at stake for the future, and the sacrifice demanded from the current

25 Parfit, *Reasons and Persons*, ch. 16.

26 This case is meant to resemble Parfit's Depletion case in *Reasons and Persons*, 362.

generation is very modest. However, this intuition is inconsistent with two very intuitively plausible principles. In other words, it does not seem possible to accept all of the following three claims:

1. For an act to be wrong, it must make someone worse-off than they would have otherwise been.
2. Being brought into existence with a life worth living does not make one worse-off than one would have been if one had never existed at all.
3. Choosing policy *B* is wrong.

Claim 1 above captures the core of the person-affecting picture of morality, as explained above, whereas claim 2 is essentially the existence-dependence claim, which we have used to get our argument for omnivorism off the ground. The problem is that if 1 and 2 are true, then choosing policy *B* is not wrong, contrary to claim 3. This is because the population that would come to exist as a result of policy *B* is not the same as the population that would come to exist as a result of policy *A*. To see this, note how transforming our way of life would lead to people working different jobs, travelling to different places, and thus meeting different people at different times—and ultimately, having different children. Consequently, as long as those who come to exist as a result of policy *B* do not have lives so bad that they are not worth living, then choosing policy *B* does not make anyone worse-off than they would have otherwise been; were we to choose differently, these people would not exist at all, and other people would exist instead. In sum, then, our defense of omnivorism commits us to denying that choosing policy *B* is wrong, contrary to our strong intuitions. Or in other words, it forces us to bite the bullet on the nonidentity problem.

The nonidentity problem is so called because most scholars have historically taken the inconsistency of the three claims sketched above to be a problem that requires a solution: we must either modify these three claims to avoid the contradiction or find a principled way to reject 1 or 2 in order to vindicate 3. To this end, one natural move is to deny 1 and appeal to policy *A* being *impersonally* better. This can then be further defined in a large variety of ways, for example by appealing to higher total or average welfare, potentially including some prioritarian weighting or only counting welfare above some critical level.[27] Alternatively, we might deny or modify 2 by adopting a different notion of harm.[28] Or to give one more example, we might think that when it comes to creating new

27 For an overview, see Greaves, "Population Axiology."

28 See, for example, Shiffrin, "Wrongful Life, Procreative Responsibility, and the Significance of Harm"; Harman, "Can We Harm and Benefit in Creating?"; and Gardner, "A Harm-Based Solution to the Non-Identity Problem."

beings, we have conditional obligations to create beings with the best possible life, even if we are not normally required to create anyone.[29]

All that said, there are also some scholars who swim against this current: most notably, David Boonin thinks that instead of seeing a problem here, we should see a sound argument, albeit with a surprising conclusion—an argument from 1 and 2 to the negation of 3.[30] Part of the reason why one might find this strategy appealing goes back to the procreation asymmetry: the impersonal approach mentioned above struggles to accommodate this asymmetry, and if one also finds none of the other strategies for responding to the nonidentity problem appealing, then one might decide that accommodating the procreation asymmetry is more important than avoiding the nonidentity problem.[31] In his paper, Podgorski sides explicitly with Boonin, claiming that the nonidentity problem need not force us to reject the case he makes for moral omnivorism.

Contra Podgorski, I believe that the cost of biting the bullet on the nonidentity problem is way too high. As I have explained, these costs include the implication that we have much less reason to fight climate change than we commonly think. But there is more: given just how contingent our identities are, this stance has a destabilizing effect on almost any putative duty to make sure future generations are better-off, whether it is about making sure that future generations can enjoy clean air, about safeguarding democratic institutions, or about aiming to secure a good start in life for our own children.[32] For me, any argument for the permissibility of eating meat that also makes it permissible to disregard the future in such a profound way simply goes too far.

Even if a reader shares my judgement, it can be objected that consumers of meat do not face a nonidentity case: instead, they face a choice between creating a given farm animal and creating no one. If this were true, then we could perhaps adopt one of the views canvassed above to explain why we should

29 For the most influential view in this vein, see Frick, "Conditional Reasons and the Procreation Asymmetry."

30 Boonin, *The Non-Identity Problem and the Ethics of Future People*.

31 For another view in this vein, see Horton, "New and Improvable Lives."

32 In "Diner's Defence," Podgorski also very briefly makes the claim that when it comes to the nonidentity problem, there is an important difference between procreation and buying meat: with the former, our intention is to create a new being, whereas with the latter, this is merely a foreseen side effect (74). He then suggests that we could bite the bullet on the nonidentity problem in the latter kind of case without doing so in the more worrying, former kind of case. Therefore, the thought goes, we need not worry about biting the bullet on the nonidentity problem in the context of eating meat. However, even if intentions were important, note that my central illustration of the nonidentity problem does not involve a choice where our intention is to create new people, and yet biting the bullet in this case still seems extremely unappealing. So this reply does not affect my argument here.

choose policy *A* in the nonidentity case while still employing the person-affecting picture to explain why buying meat is permissible.[33] However, real-life meat consumers are practically always in a nonidentity case. For one thing, they can bring into existence animals with better or worse lives depending on what kind of meat they buy. Furthermore, they can even use their money to pay for someone to raise domestic animals in a sanctuary or to conserve natural habitats. More generally, the bundle of goods we consume almost always leads to one set of beings existing rather than another. Therefore, this response does not help omnivores.

This discussion of course invites the question of what the best way to solve the nonidentity problem is. Unfortunately, the literature on the problem and the further complications it raises is so vast that I cannot engage with it here. Instead, I want to make a more modest point. What we have tried to find out is whether the kind of argument suggested by Ritchie and Stephen can be developed so that its generalization to humans seems more plausible. Despite the limitations discussed in the previous section, Podgorski's recent work helps us make progress in this direction. However, what I have shown here is that the nonidentity problem significantly curbs these ambitions: insofar as we think that we ought to help future people flourish by acting on climate change, we cannot use the argument we have been discussing to justify eating meat.[34]

33 I thank two anonymous reviewers for pushing me to clarify this point. Relatedly, Podgorski briefly mentions that the kind of intuition I rely on in my climate policy case seems to him the strongest when we must choose who to create from a set of mutually exclusive options ("The Diner's Defence," 74). He then goes on to suggest that buying meat is not like this; instead, consumers usually make a sequence of *independent* choices whereby they either create a given animal or create no one. But our consumption choices are not really independent of each other: if we spend money on a given product now, we have less money left to spend in the future, and this affects the set of options we will have. In other words, the opportunity cost of our choices means that we are always choosing from a set of mutually exclusive consumption bundles.

34 As an anonymous reviewer rightly points out, the nonidentity problem could also complicate things for the vegetarian. One possible takeaway from this section is that given the difficulties with the person-affecting picture, we should go back to the impersonal view. It could then be argued, along the lines of Singer's view, that this makes eating high-welfare meat permissible. However, there are also well-known problems with this argument: it violates the procreation asymmetry and forces us to grapple with the classic problems of population axiology, such as Parfit's "repugnant conclusion." Furthermore, I find Matheny and Chan's response to this argument convincing: if we are serious about making the world impersonally better, then animal agriculture is generally not an effective way to turn energy into welfare. However, I do not have the space here to argue for these claims. The main point is this: it might seem like Podgorski's recent work justifies omnivorism while avoiding the well-known pitfalls of the impersonal approach, but as I show, this approach leads to different problems that are just as worrying. So either way, the fact that farm

4. The Role of the Asymmetry between Producers and Consumers

In the previous section, I explained how our defense of omnivorism leads us to the nonidentity problem. In this section, I take up a more subtle issue with the argument. I believe the empirical asymmetry between producers and consumers of meat, which, as we saw in section 1, is essential for making the generalization of the argument to humans plausible, also has the effect of making the person-affecting claim at its core more appealing than it really is. Let me explain.

I assume those of us who are at least willing to entertain the starting assumption that humans and animals have the same moral status feel uneasy about the modern practice of farming animals for human consumption. Surely, we think, there must be *something* wrong about cramming sentient beings into tiny cages and killing them, at least when we do not need meat to survive and use it only to achieve momentary gustatory pleasure. Any attempt at justifying omnivorism faces the challenge of explaining away this moral anxiety.

Against this background, I believe that what makes the argument at hand initially seem so attractive is that it offers a unique explanation of this moral anxiety: we do not need to explain away this anxiety because we can instead place it firmly on the shoulders of farmers. There *is* something wrong with factory farming; it just so happens that, at the same time, consuming animal meat is permissible. We can thus use the conduct of farmers as a kind of pressure valve to relieve ourselves of any remaining moral unease we might have about our meat consumption.

Furthermore, the asymmetry between consumers and producers allows us to blame the producers in a way that many of us are already, to some extent, primed to do. This is because doing so fits together very neatly with the kind of picture that most of us, who have not worked on farms or deeply internalized the messages from animal rights activists, have about meat production. Seeing ourselves as basically decent people, we like to think that the animals we eat, as shown by the happy pigs and chickens in commercials, had good lives. Of course, we are aware that there are exceptions, but generally it is only when a rogue farmer does something terrible that this picture cracks. The picture of course suits the meat industry because it is in its interest to maintain the public perception that problems with factory farming are limited to a few rotten eggs. It is also promoted by tabloid newspapers and social media algorithms that aim to maximize clicks: news about rogue farmers doing unusually terrible things to their animals is much more shocking and thus more interesting than the

animals would not exist without us eating them does not open up an easy route to moral omnivorism. See Singer, *Practical Ethics*; Parfit, *Reasons and Persons*, 388; and Matheny and Chan, "Human Diets and Animal Welfare."

routine suffering that millions of animals on factory farms experience every day. This makes us more prone to accept the argument than we really should be.

Now, a straightforward way to test my hypothesis and to see the point I am trying to make is to imagine the asymmetry away and see what our reactions are. So let us do exactly that. Imagine that in the near future, everyone accepts our argument as it stands, and seeing that animals have legitimate complaints against many things that farmers do, we conscientiously make it illegal to farm animals. Understanding, however, that this does not mean that eating meat is inherently wrong, we replace traditional farms with fully automatized ones. To enjoy a chicken with your roast dinner, what you now need to do is go online, click one button, and wait a couple of weeks. Your click sends a signal to a production facility where a highly sophisticated 3D printer produces a fertilized egg, not involving any already existing chickens in the process. From this egg, a chicken optimized for maximal muscle growth is born, and it then lives a few weeks before its heart stops. Robots then collect the dead animal, process the meat, and deliver it to your door. The chickens suffer from a variety of health problems during their short lives due to growing so quickly, but they nevertheless have lives marginally worth living. They are also not killed by anyone—instead, just as their muscles are optimized for growth, their hearts are optimized to be just strong enough to keep the chickens alive until the desired size is reached. Furthermore, because this sophisticated process requires the chickens to be selectively bred in a very specific way, it is not possible to build these farms so that they would give the same individual animals a better life.

According to the defense of omnivorism that we have been considering, it looks like there is nothing wrong at all with this automated farming practice. Since the animals have lives not worse than nonexistence, they do not have a complaint against consumers bringing them into existence with the click of a mouse, and there is no one else who affects their fate for better or worse. Indeed, I think that if we endorse the argument for omnivorism in play here, then, as a society, moving to farms like this would be *preferable* because this would mean that we no longer accept the putatively immoral things that farmers do as mere passive onlookers. So presumably, my example should produce a neutral or even a positive reaction.

My expectation, however, is that readers feel morally troubled by this example. The primary difference between the hypothetical example and the real world is that in the hypothetical example, there are no farmers to shoulder our discomfort around eating factory-farmed meat. This suggests that when we strip away the empirical asymmetry between producers and consumers, our argument for moral omnivorism becomes less appealing, just as I suggested it would. And it seems to me that if detaching the argument from what is essentially an accidental, contingent feature of the world we currently live in makes it significantly less appealing, then this should lead us to question whether the argument is correct.

5. Farming Humans

While I find the example at the end of the previous section extremely worrying, some readers may not share my intuitions. However, in this section, I discuss a modification of that example that may lead such readers to revise their judgements. This case involves not chickens but human children. Imagine the following.

In the near future, technological progress allows us to artificially breed large numbers of children who are just like us except for a few important differences. Because the technology is not perfect, these children suffer from a variety of diseases that inevitably cause them significant pain over the course of their lives. These children also all die "naturally" at the age of seven. We farm them so that their kidneys can be harvested for lifesaving organ donations for the "regular" citizens of society. The lives of the farmed children are simple and include periods of significant pain, but they are treated as well as possible: they live in a kindergarten-like environment, where all their physical needs are met until their expiry dates arrive and they die, without us having to kill them. Thus, their lives are worth living.[35]

Even if the argument we have been considering avoids Salt's charge that it licenses holding slaves or parents treating their children however they wish, nothing in the argument blocks the implication that farming children for their organs along the lines of the above example is permissible. The farmed children have no complaint against consumers buying organs from farms like this because the children's existence depends on the demand for organs, and this existence is not worse than nonexistence. Even more importantly, however, the permissibility here is not limited to the consumers because the children are (by stipulation) treated as well as possible, and they die without us having to slaughter them. This means that the producers working on these farms are not wronging the children either. In fact, we can just imagine the producers away, just like we did with the automated chicken farm: perhaps the children come to term in artificial wombs, get taken care of by humanoid robots, and

35 Interestingly, while this example is fictional, note that a kidney from a genetically modified pig has already been used in a human recipient. For a recent example of this procedure, see Yousif, "Pig Kidney Transplant Patient Leaves Hospital." For a discussion of the ethical issues that this practice raises, including arguments from the nonidentity problem, see Casal and Williams, "Human iPSC-Chimera Xenotransplantation and the Non-Identity Problem."

finally have their organs processed in a fully automated production line.[36] So our argument for omnivorism implies that there is nothing wrong at all with this practice. Yet I take it that the thought of having an organ from one of these farms strikes us as extremely morally troubling.

In fact, we can go even further. Note that from a strictly person-affecting point of view, these human kidney farms are much better than regular animal farms, as the already existing people receiving the kidneys get a benefit that is much more significant than mere gustatory pleasure. Indeed, on the view we are considering, if we imagine that building and operating these farms turns out to be cheap and easy, then it starts to look like we may even be *obliged* to build them. For if we do not, then already existing people who need kidney transplants to avoid death could have a weighty moral complaint against the rest of us on the basis that we have gratuitously failed to help them by refusing to engage in a practice that saves the lives of many people while generating complaints from no one.

This result strikes me as absurd. It shows that even the most developed version of the argument for omnivorism that we have considered has very counterintuitive implications when applied to humans, even when we leave the nonidentity problem aside.

Even though I believe that the example I have just described gives us a weighty reason to reject the argument for omnivorism that we have been considering, there is a natural objection to make here: What about the case of Theodore? In that case, the argument for omnivorism that we are considering seems to have no troubling implications, even when applied to humans, whereas the case I just presented seems very troubling indeed. This conflict of intuitions, it might be thought, calls for an explanation from vegetarians.[37] I think there are two plausible responses here.

First, while I previously took the reaction Podgorski prescribed to the case involving Theodore as given, I think there is in fact space to push back. The intuition that buying the chocolate and thereby creating Theodore is permissible is likely to be skewed by a general resistance to what we perceive as overdemandingness. The idea that something as ordinary as buying a chocolate bar might be impermissible due to some complicated causal chain leading to a given person being born might well strike us as too demanding: it feels unreasonable to be subjected to moral criticism just because a seemingly innocuous everyday action happens to have certain consequences for far-off future people. But we might, on reflection, nevertheless decide that there are legitimate moral

36 I thank an anonymous reviewer for prompting me to make this link clearer.

37 I thank an anonymous reviewer for pushing me to clarify my view regarding this objection.

demands of this type: perhaps we should aim to adjust our consumption habits, career choices, and activism so that these tend towards better-off beings existing in the future, at least when doing so is not too costly. Many of us already accept something like this in relation to climate change: we might, for example, buy secondhand clothes or travel by train instead of flying. So perhaps we can reconcile our judgements by accepting that depending on the details, creating Theodore may not be permissible after all.

Note that this move becomes even more appealing when we consider just how strong the intuition against human farming is: if we have to give up our intuitions in either the case of Theodore or the case considered here, it seems better to give up the intuition regarding Theodore. Furthermore, if we accept a view on the nonidentity problem that, in one way or another, sets a high bar for the permissibility of creating new people, then such a view naturally goes together with this response. To put my cards on the table: I think this is most likely the right way to go. However, defending this view requires taking a stance on various issues in population ethics, and doing so is not possible in this article. So it is worth considering another plausible way for vegetarians to reconcile our judgements.

The second option is to retain our original reactions to both the human farming case and the case involving Theodore by appealing to some morally relevant difference between them. One natural candidate for such a difference is that creating Theodore is a side effect of buying the chocolate bar, whereas by farming humans for organs, we use them as means to our ends.[38] Interestingly, Podgorski in fact grants that something like this reasoning may still pose an objection to omnivorism, even if we accept his reasoning.[39] However, he maintains that such an objection is much weaker than the harm-based objections usually raised against eating meat.

Personally, I am not sure how much weight to put on objections to omnivorism that rely on something like the idea that we should not use other beings as mere means to our ends. Nevertheless, if this is what distinguishes

38 For one version of this idea in the context of creating people, see Kavka, "The Paradox of Future Individuals."

39 Podgorski discusses the related but different idea that it may be wrong to benefit from wrongdoing because he operates under the assumption that producers of meat wrong animals, even though consumers do not ("The Diner's Defence," 75). However, imagining farms to be automated in the way I have described makes this idea difficult to employ, which is why I talk about using others as means. Alternatively, we might try to appeal to the doctrine of double effect or some version of the ideal of relational equality: perhaps the problem with farming is that it brings about a world where farmed humans have a status of second-class citizens, and there could be something objectionable about this even if these people are not straightforwardly harmed by their existence.

the case involving Theodore from the case I presented, then clearly what we have here is a very weighty moral consideration. After all, this consideration makes a huge difference in the intuitions that we have with regard to the two cases. If we remain consistent on our commitment to the equal moral status of humans and farm animals, then it follows that this consideration also poses a very serious objection to omnivorism—contrary to what Podgorski claims. In sum then, even if we hold on to our initial judgement regarding the case involving Theodore, eating meat remains clearly impermissible.

There is one more natural thought that I believe some readers may have in response to the human farming case. This is the thought that even if there are significant moral limits to how we may treat animals, the moral status of farm animals must be different from that of humans after all. It could be that the difference between the moral status of humans and animals is something like this: while neither can be harmed, humans should also not be treated as mere means. If using others as mere means is the key difference between the human farming case and the case involving Theodore, then this view would entail that human farming is impermissible, whereas creating Theodore and buying meat, even from an automated farm, are both permissible. I expect that some readers find this combination of results appealing.[40]

However, note that even if a view like this is easier to defend than the extreme view that farm animals do not morally matter at all, it still faces significant costs. For one thing, if omnivores take this route, then they abandon Podgorski's ambitious project of justifying omnivorism even while assuming equal moral status. But more importantly, views like this face all the classic problems of justifying differential moral status for humans and animals: we need to explain exactly what gives rise to the additional moral constraints on our treatment of humans and to do this in a way that deals fairly with cases on the margins, including humans with limited cognitive capacities and animals with capacities more sophisticated than that.[41] While it may be true that the argument for omnivorism that I have considered in this article remains relevant, taking on this project would amount to accepting that the philosophical heavy lifting happens elsewhere.

Finally, note that even if we manage to justify the differential moral status of humans and animals along the lines discussed above and accept that this deals with the human farming case, the resulting defense of omnivorism still commits us to biting the bullet on the nonidentity problem. Choosing what

40 I thank an anonymous reviewer for prompting me to think about this possibility.

41 For a helpful statement and clarification of these issues from a vegetarian perspective, see Horta, "The Scope of the Argument from Species Overlap."

I call policy *B*, which leads to a future with much lower levels of welfare than the alternative policy, does not generate complaints from future people, nor does it treat them as means—rather, their existence is the side effect of our lifestyle choices. Therefore, even if we concede all this ground to omnivores, the argument that eating meat is permissible because the animals we eat would not exist otherwise comes with a hefty theoretical cost.

6. CONCLUSION

Let us recap. I began by reconstructing what I take to be the best version of the well-known argument that eating meat is permissible because the animals would not exist if we did not farm them for our consumption. I argued that even the most sophisticated version of this argument is limited in two important ways: first, it cannot justify the consumption of all animal products, including dairy; and second, it still implies that we should stop animals from being harmed by producers if we can.

After reconstructing the argument and pointing out its limitations, I raised three objections. First, I argued that the nonidentity problem poses a significant issue for the argument. Second, I argued that the asymmetry between producers and consumers that seems necessary to make the generalization of the argument to humans plausible also has the effect of giving it undue credibility. Finally, I argued that even the most developed version of the argument is still very difficult to accept when generalized to humans.

All of my objections give rise to some complications. Not all scholars share my judgement on the nonidentity problem; the automated chicken farm may generate different intuitions; and the difference between the intuitions we have regarding human farms and the case involving Theodore may strike us as puzzling. Nevertheless, my view is that taken together, these objections give us a decisive reason to reject the argument for omnivorism that we started with. It could still be the case that eating meat is permissible. But if it is, this is not because the animals we eat would not have existed without us farming them.[42]

University College London
karri.heikkinen.20@ucl.ac.uk

42 For comments on this article, I thank Joe Horton, Theron Pummer, Ben Sorgiovanni, Euan Allison, and two anonymous reviewers. For helpful discussion, I also thank participants at the University College London PhD work-in-progress seminar, the 2023 University College London Graduate Conference in Philosophy, and the 2025 Festival of Ideas at Universitat Pompeu Fabra. Research for this article was financially supported by the London

References

Boonin, David. *The Non-Identity Problem and the Ethics of Future People*. Oxford University Press, 2014.

Casal, Paula, and Andrew Williams. "Human iPSC-Chimera Xenotransplantation and the Non-Identity Problem." *Journal of Clinical Medicine* 8, no. 95 (2019): 1–13.

Delon, Nicolas. "Replaceability Argument in the Ethics of Animal Husbandry." In *Encyclopedia of Food and Agricultural Ethics*, edited by David M. Kaplan. Springer, 2019.

De Vries, Albert, and Marcos Inácio Marcondes. "Review: Overview of Factors Affecting Productive Lifespan of Dairy Cows." *Animal* 14, no. 1 (2020): 155–64.

Frick, Johann. "Conditional Reasons and the Procreation Asymmetry." *Philosophical Perspectives* 34, no. 1 (2020): 53–87.

Gardner, Molly. "A Harm-Based Solution to the Non-Identity Problem." *Ergo* 2, no. 17 (2015): 427–44.

Goldhill, Olivia. "An Oxford Philosopher's Moral Crisis Can Help Us Learn to Question Our Instincts." *Quartz*, October 15, 2017. https://qz.com/1102616/an-oxford-philosophers-moral-crisis-can-help-us-learn-to-question-our-instincts.

Greaves, Hilary. "Population Axiology." *Philosophy Compass* 12, no. 11 (2017): e12442.

Hare, R. M. "Why I Am Only a Demi-Vegetarian." In *Essays on Bioethics*. Oxford University Press, 1999.

Harman, Elizabeth. "Can We Harm and Benefit in Creating?" *Philosophical Perspectives* 18, no. 1 (2004): 89–113.

Hedden, Brian. "Consequentialism and Collective Action." *Ethics* 130, no. 4 (2020): 530–54.

Horta, Oscar. "The Scope of the Argument from Species Overlap." *Journal of Applied Philosophy* 31, no. 2 (2014): 142–54.

Horton, Joe. "New and Improvable Lives." *Journal of Philosophy* 118, no. 9 (2021): 486–503.

John, Tyler, and Jeff Sebo. "Consequentialism and Nonhuman Animals." In *Oxford Handbook of Consequentialism*, edited by Douglas W. Portmore. Oxford University Press, 2020.

Kagan, Shelly. "Do I Make a Difference?" *Philosophy and Public Affairs* 39, no. 2

Arts and Humanities Partnership and the Forethought Foundation, both of which are gratefully acknowledged.

(2011): 105–41.
Kavka, Gregory. "The Paradox of Future Individuals." *Philosophy and Public Affairs* 11, no. 2 (1982): 93–112.
Matheny, Gaverick, and Kai M. A. Chan. "Human Diets and Animal Welfare: The Illogic of the Larder." *Journal of Agricultural and Environmental Ethics* 18, no. 6 (2006): 579–94.
McMahan, Jeff. "Causing People to Exist and Saving People's Lives." *Journal of Ethics* 17, nos. 1–2 (2013): 5–35.
———. "Eating Animals the Nice Way." *Daedalus* 137, no. 1 (2008): 66–76.
Nefsky, Julia. "Collective Harm and the Inefficacy Problem." *Philosophy Compass* 14, no. 4 (2019): e12587.
———. "How You Can Help, Without Making a Difference." *Philosophical Studies* 174, no. 11 (2017): 2743–67.
Norcross, Alastair. "Puppies, Pigs, and People: Eating Meat and Marginal Cases." *Philosophical Perspectives* 18, no. 1 (2004): 229–45.
Parfit, Derek. *Reasons and Persons*. Clarendon Press, 1984.
Podgorski, Abelard. "The Diner's Defence: Producers, Consumers, and the Benefits of Existence." *Australasian Journal of Philosophy* 98, no. 1 (2020): 64–77.
Regan, Tom. *The Case for Animal Rights*. University of California Press, 1983.
Ritchie, David George. *Natural Rights: A Criticism of Some Political and Ethical Conceptions*. Swan Sonnenschein, 1895.
Salt, Henry S. *The Humanities of Diet*. Vegetarian Society, 1914.
Shiffrin, Seana. "Wrongful Life, Procreative Responsibility, and the Significance of Harm." *Legal Theory* 5, no. 2 (1999): 117–48.
Singer, Peter. *Animal Liberation*. Harper Collins, 1975.
———. *Practical Ethics*. 2nd ed. Cambridge University Press, 1993.
Stephen, Leslie. *Social Rights and Duties: Addresses to Ethical Societies*, vol. 1. Swan Sonnenschein, 1886.
Višak, Tatjana. *Killing Happy Animals: Explorations in Utilitarian Ethics*. Palgrave Macmillan, 2013.
Yousif, Nadine. "Pig Kidney Transplant Patient Leaves Hospital." BBC News, April 4, 2024. https://www.bbc.com/news/world-us-canada-68710229.
Zuidema, Dalen, Karl Kerns, and Peter Sutovsky. "An Exploration of Current and Perspective Semen Analysis and Sperm Selection for Livestock Artificial Insemination." *Animals* 11, no. 12 (2021): 3563.

Journal of Ethics and Social Philosophy
vol. 30, no. 6 · october 2025
https://doi.org/10.26556/jesp.v30i6.4419

THE PRACTICE ACCOUNT OF POLITICAL AUTHORITY

Fabian Wendt

The most fundamental problem of political philosophy is to explain the authority of the state. This article presents a novel account: the practice account of political authority. The practice account belongs to the family of natural duty accounts, but in contrast to other such accounts, it highlights the role that rules of conventional practices can (and must) play in explaining political authority. The idea is that we have a natural right to participate in justifiable conventional practices that secure basic justice and peace, as well as a natural duty to respect this same right in others; that duty explains why the rules of such practices can confer political authority on the state.[1] Other natural duty accounts, the article shows, cannot explain political authority because they fail to provide a working mechanism that links people's natural duties to the authority of the state.

The article proceeds as follows: it introduces the concept of political authority (section 1), lays out the general form of natural duty accounts of political authority (section 2), and explains why natural duty accounts as they have been developed so far cannot explain political authority (section 3). It then introduces the practice account of political authority as an alternative natural duty account that explains political authority (section 4), and it responds to objections (section 5).

1. Political Authority

Let me start by clarifying what I mean by political authority, i.e., the *explanandum* of this article. Political authority is a bundle of rights held by the state.[2] I take the core of political authority to be the Hohfeldian power to impose

1 The basic idea (or something more or less close to it) can be found in Wendt, "Political Authority and the Minimal State," 115–16, although I dismissed it back then.

2 On Hohfeld's distinction between different types of rights (claim rights, liberty rights (or privileges), powers, and immunities) and their correlates (duties, no-rights, liabilities, and disabilities), see Hohfeld, *Fundamental Legal Conceptions As Applied in Judicial Reasoning*. See also Kramer, "Rights Without Trimmings"; and Wenar, "The Nature of Rights."

duties.[3] Political authority arguably also includes liberty rights to create and enforce laws, claim rights against interference with the creation and enforcement of laws, and perhaps more.[4] But the power to impose duties is my focus in this article. In Hohfeldian terms, the correlate of the state's power to impose duties on the side of citizens is not a duty to obey the law; it is a liability to have duties imposed by the state. (A duty to obey the law would be the correlate of the state's claim right to be obeyed.)

Political authority is a set of moral rights, not legal rights. That states have legal liberty rights to create and enforce laws and the legal power to impose legal duties is uncontroversial. What is controversial is whether states have the corresponding *moral* liberty rights to create and enforce laws and the *moral* power to impose moral duties (by exercising their legal power to impose legal duties). While it is the state that has political authority, more specific state institutions (parliaments, governments, the police, etc.) have more specific rights that form a part of the state's political authority. These institutions in turn come with positions that are to be held by particular persons. In the end, it is particular persons who have the rights of a prime minister, member of parliament, or police officer, for example.

Political authority, as I would like to conceive it, is *general*. It holds vis-à-vis everyone on the state's territory and all citizens. Political authority is not a patchwork notion that holds only vis-à-vis some people on the territory and some citizens, or for different citizens to different degrees.

Another important point is that the power of the state to impose duties is *content independent*. By enacting laws, states impose duties independently of the law's content. When the state says that you are to drive on the right side of the street, you have a duty to drive on the right side of the street; when it says that you are to drive on the left side, you have a duty to drive on the left side. Content independence does not imply that there are no moral limits regarding the content. Rather, within the range of what is morally permissible, agents

3 Cf. Copp, "The Idea of a Legitimate State"; Applbaum, "Legitimacy Without the Duty to Obey"; Perry, "Political Authority and Political Obligation"; and Enoch, "Authority and Reason-Giving." On Applbaum's view, more precisely, "legitimate authority has the moral power to author legal, institutional, or conventional rights and duties, powers, and liabilities, and create social facts and mechanisms of coordination that change the legal, institutional, and conventional situation or status of subjects" ("Legitimacy Without the Duty to Obey," 221), and whether the exercise of this moral power gives rise to moral duties or changes the subjects' moral situation in other ways is a separate question.

4 It probably also includes powers to alter people's rights in ways than cannot adequately be described as an imposition of duties. For example, a state could deprive people of their power to sell certain goods to sixteen-year-olds (Brinkmann, *An Instrumentalist Theory of Political Legitimacy*, 32).

with powers to impose duties can create duties independently from the duties' content. Accordingly, political authority is *not unlimited*. It is the moral right to make and enforce law within moral limits. To ascribe political authority and thus a power to impose duties to the state is perfectly compatible with a belief in limited government. As far as possible, I leave open here what these moral limits are.

I think that states need political authority—in particular, the power to impose duties on everyone on the territory and all citizens—if they are to count as legitimate.[5] This motivates my search for an account that can explain how states can come to have that power. But legitimacy is not my topic here. Even if states did not need political authority to be legitimate, it would be of interest, of course, if and when they have political authority.

2. NATURAL DUTY ACCOUNTS

Because political authority is general—it applies to all citizens and to everyone on the state's territory—voluntarist accounts have difficulties vindicating the state's political authority: if authority is to be based on voluntary consent, as in consent accounts, then states do not have authority over those who do not consent; if authority is to be based on the voluntary acceptance of benefits provided by the state, as in fair play accounts, then states do not have authority over those who do not voluntarily accept those benefits.[6] Likewise, associativist accounts have to deny that states have authority over those who do not associate with the political community in the required way.

The most promising accounts of political authority are therefore natural duty accounts. Natural duties are duties that persons have not in virtue

5 The main reason is that states must have the moral right to do what they are doing if they are to count as legitimate, and since states exercise a legal power to impose duties, they must have the corresponding moral power. Cf. Wendt, "Justice and Political Authority in Left-Libertarianism," 328–32, and *Authority*, 101–3.

6 Klosko's fair play account assumes that everyone *would* voluntarily accept indispensable goods *if* this were necessary to receive them (*The Principle of Fairness and Political Obligation*, ch. 2). Simmons suggests that Klosko's account ultimately relies on an unacknowledged natural duty to make indispensable goods available to everyone ("The Duty to Obey and Our Natural Moral Duties," 189–90; for a reply, see Klosko, "Fair Play, Reciprocity, and Natural Duties of Justice"). Similarly, in Kavka's hypothetical consent account (*Hobbesian Moral and Political Theory*, ch. 10) and Estlund's normative consent account (*Democratic Authority*, ch. 7), it is arguably a natural duty that is ultimately to justify political authority, while the idea that people *would* or *ought to* give their consent to political authority is at best an intermediate step. Cf. Sreenivasan, "Oh, but You Should Have"; Koltonski, "Normative Consent and Authority"; Wendt, *Authority*, 29–33; and Zhu, "Content-Independence and Natural-Duty Theories of Political Obligation," 69–71.

of voluntary acts or contingent memberships but in virtue of their humanity. Jonathan Quong suggests the following principle to connect natural duties and authority:

> One way to establish that a person has legitimate authority over another person involves showing that the alleged subject is likely better to fulfil the duties of justice he is under if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to directly fulfil the duties he is under himself.[7]

The principle is inspired by Joseph Raz's normal justification thesis.[8] But while Raz's thesis claims that authority arises when by complying with someone's directives, people better comply with reasons that independently apply to them, Quong's principle is narrower: it claims that authority arises when by complying with someone's directives, people better fulfill their natural duties.

Quong illustrates his principle with an example in which a doctor and a person without medical expertise arrive at the scene of an accident.[9] Under these circumstances, the best way for the latter to fulfill their natural duty to help the victims is to follow the directives of the doctor. The doctor thus acquires a power to impose duties, and the other person incurs a liability to have duties imposed by the doctor. The explanation of political authority follows a similar path: people have certain natural duties, they best fulfill these duties by incurring a liability to have duties imposed by the state, and the state accordingly acquires a power to impose duties on them.

In one respect, the doctor example is misleading: people better comply with their natural duties by following the doctor's directives because the doctor has greater medical expertise. In the case of the state, it is doubtful that we can expect the state's representatives to have greater expertise in lawmaking; at least we do not have reliable means of identifying experts in this realm, while there are such means for medical expertise.[10] The reason the state is in the position to help people better fulfill their natural duties has arguably less to do with its expertise and more with the benefits it can provide as a central coercive agency that provides and enforces binding rules for everyone. It is the state's structural features, not epistemic ones. But what could be the underlying natural duty that grounds the state's political authority? Let me briefly introduce three natural

7 Quong, *Liberalism Without Perfection*, 128.

8 Raz, *The Morality of Freedom*, 53.

9 Quong, *Liberalism Without Perfection*, 127. See also Estlund, *Democratic Authority*, 124.

10 Zhu, "Content-Independence and Natural-Duty Theories of Political Obligation," 68–69.

duties that have been proposed: the natural duty of justice, the natural duty of rescue, and the natural duty not to harm others.

John Rawls assumes that persons have a natural duty of justice, which has two parts: "First, we are to comply with and to do our share in just institutions when they exist and apply to us; and second, we are to assist in the establishment of just arrangements when they do not exist, at least when this can be done with little cost to ourselves."[11] The first part is what matters for our context, since we are interested in how and why existing states have political authority.

For Rawls, this natural duty is to account for people's political obligations, in particular the obligation to comply with the law; but when we add the assumption that people do not have a right "not to be coerced to do what we have an obligation of justice to do," then we have an argument why just states at least have a liberty right to enact and coercively enforce laws.[12] This still does not amount to political authority, though, since political authority contains more than that liberty right. It also includes a power to impose duties. But Quong's principle is to help take the final step to political authority: in order to fulfill their natural duty of justice to comply with and to do their share in just institutions when they exist and apply to them, people incur a liability to have duties imposed by the state, and the state gets a power to impose duties on them.

Since hardly any state is perfectly just, one may think that Rawls's duty of justice is a nonstarter: it does not apply to any state that actually exists. But one need not read the natural duty of justice as a duty to comply only with perfectly just institutions—one can (and should) read it as a duty to comply with institutions that are "reasonably just."[13]

And another modification seems necessary. There are potentially many reasonably just institutions that "apply to us" in some sense, and we are certainly not under a duty to comply with all of them. The NAACP may be said to "apply" to all people of color, and it may further the cause of justice, but it certainly does not have a power to impose duties.[14] A more plausible duty of justice requires us to comply with institutions that apply to us and are *necessary* to establish

11 Rawls, *A Theory of Justice*, 334 (cf. 115). Other philosophers conceptualize the natural duty of justice differently (e.g., Buchanan, "Political Legitimacy and Democracy," 703; Lefkowitz, "Simmons' Critique of Natural Duty Approaches to the Duty to Obey the Law," 10–11; Christiano, *The Constitution of Equality*, 249; Stilz, *Liberal Loyalty*, 54; and Layman, "Rights, Respect, and Equality," 98, 108, 111–12). I am not able to go into the details of these accounts; I think the problem of natural duty accounts to explain political authority that I describe in the next section applies to all of them.

12 Buchanan, "Political Legitimacy and Democracy," 703.

13 Quong, *Liberalism Without Perfection*, 132–35.

14 Simmons, "The Duty to Obey and our Natural Moral Duties," 162.

and maintain justice because compliance with those institutions is necessary to discharge our natural duty of justice to a sufficient degree. Institutions that apply to us and are necessary to establish and maintain justice thus acquire a power to impose duties on us, and we incur a liability to have duties imposed by them. Relatedly, institutions do not have authority if people "likely better" fulfill their natural duties of justice by submitting to the authority of the state, as Quong puts it in the quote above; rather, institutions have authority if people *must* submit to their authority if they are to fulfill their natural duties of justice.

Christopher Wellman appeals to a natural duty of rescue, not a natural duty of justice. The natural duty of rescue is a natural duty to save others from peril when we can do so without unreasonable costs for ourselves and when our help is necessary.[15] This duty applies to accidents as in the example above; but according to Wellman, the Hobbesian state of nature is also something we have a natural duty to rescue each other from. Since the state is necessary to keep us out of the Hobbesian state of nature, we discharge our duty to rescue each other from the state of nature (at least) by letting the state do its job, and the state has the liberty right to exercise the coercion that is necessary to save us from the Hobbesian state of nature.[16]

This liberty right does not amount to political authority since, again, political authority also contains the power to impose duties. But while Wellman does not do so (more on this below), one could try to link the natural duty of rescue to the state's power to impose duties: like the only way to sufficiently discharge one's natural duty of rescue is to follow the directives of the doctor in the above example, the only way to sufficiently discharge one's natural duty of rescue is to follow the directives of the state, and that is why the state has the authority to impose duties on persons by making law and, on the flip side, why persons have a liability to have duties imposed by the state.

Massimo Renzo's account works with neither a natural duty of justice, nor a natural duty of rescue; rather, he relies on a natural duty not to harm others.[17] Unlike the duty of justice and the duty of rescue, the duty not to harm is conceptualized as a purely negative duty. Renzo does intend that duty to explain proper political authority, including the power to impose duties. Because the state is necessary to save us from the perils of the state of nature, one harms

15 Wellman, "Liberalism, Samaritanism, and Political Legitimacy," 215, "Toward a Liberal Theory of Political Obligation," 744–45, and "Samaritanism and the Duty to Obey the Law," 21–23.

16 Wellman, "Liberalism, Samaritanism, and Political Legitimacy," 216–19, 223, "Toward a Liberal Theory of Political Obligation," 745–48, and "Samaritanism and the Duty to Obey the Law," 23.

17 Renzo, "State Legitimacy and Self-Defence."

others—one poses an unjust threat to those who are nearby—by not subjugating to the authority of the state.[18] To discharge one's natural duty not to harm others, one therefore incurs a liability to have duties imposed by the state, and the state gets the power to impose duties.

3. The Failure to Explain Political Authority

In the previous section, I explained how natural duty accounts of political authority are supposed to work, and I introduced the natural duty of justice, natural duty of rescue, and natural duty not to harm others as three natural duties that have been invoked by different philosophers. As we saw, for states to have political authority, it must be the case that because the state is necessary to provide justice and peace, people cannot discharge their natural duties (of justice or of rescue or not to harm others) without a liability to have duties imposed by the state. In other words, natural duty accounts of political authority must claim that "in order for a state to be able to carry out its proper functions . . . , it must have authority over everyone living on its territory."[19]

But it is actually doubtful that this is the case.[20] What the state needs to properly function and provide justice and peace is, at best, to be able to impose duties on a critical number of people—but not on everyone. As long as a sufficient number of people obey the law, the rest may discharge their natural duties to a sufficient degree by not disrupting the operations of the state and by complying with the law whenever they deem appropriate. The proposed mechanism that is to generate political authority does not work.

In response, Renzo argues that instability would result if the state did not have a power to impose duties on everyone, but only a power to impose duties on some:

> The point is that in this scenario nobody could rely on the fact that independents [who are not subject to the state's authority] would be acting in accordance with what the law says (even when they would have moral or prudential reasons to do so), for everyone would know

18 Renzo, "State Legitimacy and Self-Defence," 585–86.

19 Renzo, "State Legitimacy and Self-Defence," 586.

20 Cf. Buchanan, "Political Legitimacy and Democracy," 695–96; Simmons, "The Duty to Obey and Our Natural Moral Duties," 168–69, 175–77, 182, 187–88, and *Boundaries of Authority*, 84; Adams, "Institutional Legitimacy," 98–100; Zhu, "Farewell to Political Obligation," 459–60, and "Content-Independence and Natural-Duty Theories of Political Obligation," 72–73, 75–76.

> that how independents are going to act will ultimately depend on their fallible individual judgement.[21]

In the end, though, he concedes that this instability might not be severe enough to undermine the state's ability to perform its vitally important functions. He thus concedes that states may have authority only over the number of people that is needed for the state to function.[22] This of course is not proper political authority.

Wellman does not claim that natural duties could ground political authority. For him, they underly only the state's liberty right to do what is necessary to keep us out of the Hobbesian state of nature. Not even political obligations can be fully explained in terms of natural duties, according to him, because the state does not need universal compliance to do its job. Wellman appeals to a principle of fair play to account for people's political obligations: it would be unfair to claim the discretion to decide how to discharge one's natural duty of rescue for oneself when we could not escape the Hobbesian state of nature if everyone kept this discretion.[23] Could one make a similar argument to account for the state's power to impose duties and thus its political authority? The argument could be that to properly function, the state needs to be able to impose duties on *some*, and it would be unfair if only some would have a liability to have duties imposed by the state, while others are off the hook.

But in reply, one can doubt that the state actually needs a power to impose duties on *anyone* to properly do its job. What the state needs in order to properly function is that a sufficient number of people comply with the law, but this does not require a liability to have duties imposed by the state. People can comply with the law for all kinds of reasons: they can comply to avoid punishment, they can comply because they are lazy, or they can comply for moral reasons (for example, because they judge the law to be congruent with moral demands or because they value stability and do not want to disrupt the state's operations). People may thus be perfectly able to sufficiently discharge their natural duties by complying with the law for any of these reasons, without a liability to have duties imposed by the state. And if the state does not need a power to impose duties on anyone in order to properly function, then no issues of fairness and unfairness arise either.

21 Renzo, "State Legitimacy and Self-Defence," 594.

22 Renzo, "State Legitimacy and Self-Defence," 595.

23 Wellman, "Toward a Liberal Theory of Political Obligation," 749–51, and "Samaritanism and the Duty to Obey the Law," ch. 2.

4. The Practice Account

Natural duty accounts as they have been developed so far cannot explain political authority, as we saw in the previous section: people can sufficiently discharge their natural duties of justice or duties of rescue or duties not to harm others without incurring a liability to have duties imposed by the state.

Understanding the role that rules of conventional practices can play in generating authority is, I think, crucial to develop a more potent natural duty account. That rules in cooperative schemes can create authority relations has been suggested by H. L. A. Hart and, more recently, Justin Tosi.[24] Hart writes:

> [When] a number of persons conduct any joint enterprise according to rules and thus restrict their liberty, those who have submitted to these restrictions when required have a right to a similar submission from those who have benefited by their submission. The rules may provide that officials should have authority to enforce obedience and make further rules, and this will create a structure of legal rights and duties, but the moral obligation to obey the rules in such circumstances is *due to* the co-operating members of the society, and they have the correlative moral right to obedience.[25]

Hart and Tosi take the rules of cooperative schemes to create authority relations against the background of fair play obligations to comply with the rules. They do so, explains Tosi, because they specify the terms of cooperation.[26] As indicated above, I do not think everyone has fair play obligations, because some will not voluntarily accept the benefits provided by the state, and so I do not think a fair play account can explain political authority if political authority is to apply to everyone on the territory and to all citizens.

The idea, then, is to combine Hart's and Tosi's insight about the role of rules with a natural duty account. Because of the role that conventional practices and their rules play in generating authority on this account, I call it *the practice account of political authority*, but I still take it to be a member of the family of natural duty accounts of political authority.[27]

24 Hart, "Are There Any Natural Rights?"; and Tosi, "The Possibility of a Fair Play Account of Legitimacy."

25 Hart, "Are There Any Natural Rights?" 185.

26 Tosi, "The Possibility of a Fair Play Account of Legitimacy," 94–95.

27 Melenovsky distinguishes five types of reasons that, depending on the context, can justify following the rules of conventional practices ("The Reasons to Follow Conventional Practices"). My proposal here invokes a reason that "specifies how to respect a preexisting right, responsibility, or duty" (720).

From the perspective of a natural duty account, rules that confer authority are not taken to specify the terms of a cooperative scheme that cooperators have fair play obligations to comply with. Instead, the rules that confer authority are conceived as part of a conventional scheme that people have a natural duty to respect. More precisely, I picture the following argument:

1. Persons have a natural right to participate in justifiable conventional practices that secure basic justice and peace (in short, a "natural right of basic justice and peace"). Among other things, this natural right includes a moral power to acquire new moral claim rights, liberty rights, and powers in line with the rules of such practices.
2. Persons have a natural duty to respect other persons' natural right to participate in justifiable conventional practices that secure basic justice and peace (in short, a "natural duty of basic justice and peace"). Among other things, this natural duty includes a moral liability to have one's claim rights and liberty rights changed when others acquire and exercise new moral claim rights, liberty rights, and powers in line with the rules of such practices.
3. The rules of justifiable conventional practices that secure basic justice and peace can confer authority on state institutions, including a content-independent power to impose duties on everyone on the territory and all citizens.
4. If the rules of a justifiable conventional practice that secures basic justice and peace confer a content-independent power on state institutions to impose duties on everyone on the territory and all citizens, then the natural right of basic justice and peace allows persons to obtain positions within these institutions in accordance with the rules of the practice and to thereby acquire a content-independent power to impose duties on everyone on the territory and all citizens. (From 1 and 3)
5. If the rules of a justifiable conventional practice that secures basic justice and peace confer a content-independent power on state institutions to impose duties on everyone on the territory and all citizens, then the natural duty of basic justice and peace puts all citizens and everyone on the state's territory under a liability to have duties imposed on them by state institutions. (From 2 and 3)

Before commenting on each step in the argument, let me say a few words on how I understand the terms 'conventional practice', 'peace', and 'basic justice'. This article is not to contribute to debates about what conventions, peace, and justice are, and the argument of the article is supposed to work for a variety of conceptions of what they are. But roughly, a conventional practice can be understood as a set of primary rules that is to guide people's behavior, sometimes combined with secondary rules that determine how primary rules can be changed, and these rules must be widely accepted and internalized within some social group. For the purposes of this article, one should think of legal systems and social norms as relevant conventional practices. Peace, on the most commonsense conception that I endorse here, is the relatively stable absence of violence in a particular geographical region. By basic justice, I mean those elements of justice that are uncontroversial—that everyone's physical integrity and personal property are to be respected, that conflicts are to be solved in fair ways, that punishment is to be proportional to the crime, and so on. Conventional practices that secure and maintain peace and basic justice can involve lawmaking institutions like parliaments combined with state police, military, and courts; but they can also center around nonstate equivalents—for example, private protection companies or religious courts.

Let me now comment on the separate steps in the argument. The first premise introduces the natural right of basic justice and peace, which is a natural right to participate in justifiable conventional practices that secure basic justice and peace. This natural right is to be conceived as a bundle of rights that, among other things, comprises a Hohfeldian claim right against being excluded from such practices, a liberty right to participate in them, and a power to acquire new claim rights, liberty rights, and powers in ways delineated by those practices. The latter power is what matters for our context. It means, for example, that one can acquire the rights associated with citizenship, as determined by the rules of the practice, and that one can become a member of parliament with the rights this position involves, when this is something that the practice allows.

All these rights are to be understood as *moral* rights. Conventional practices of course allow people to acquire legal (or otherwise conventional) claim rights, liberty rights, and powers, first of all. But that persons have a natural right to acquire rights in line with such practices means that these rights are morally validated: people who acquire them also acquire the corresponding moral rights. Compare a different case. In soccer games, coaches have the power to substitute players, which implies the power to impose duties on players to leave the field. This power and these duties are, first of all, purely conventional. But if soccer coaches have the moral right to do what they do, then they also have the corresponding moral power to substitute players and thereby to impose duties on players to leave the field (or, if you prefer, the moral liberty right to exercise their conventional power to impose duties on players to leave the field). And if coaches have this moral power (or liberty right), then likewise, players incur not just a conventional but also a moral duty to leave the field when coaches tell them to.

Importantly, the natural right of basic justice and peace is a right to participate in *justifiable* practices. To be justifiable, a practice does not have to be morally perfect. It has to meet some threshold of moral justifiability, and whether a practice actually secures and maintains peace and basic justice to a sufficient degree is part of what determines that threshold. Beyond that, I want to leave open what the justifiability standard amounts to and thus keep the practice account compatible with as many views in moral and political philosophy as possible. The point of this article is not to defend a particular standard of justifiability but to propose a mechanism that explains how states can acquire political authority against the background of *some* standard of justifiability.

Two more points are worth mentioning. One is that what counts as justifiable is probably relative to what is feasible given the circumstances of a particular time and place. The other is that one can distinguish between the justifiability of a practice *tout court* and the justifiability of a practice relative to a particular individual. The natural right of basic justice and peace appeals to justifiability *tout court*. It is possible of course to hold that a practice is justifiable *tout court* only if it is justifiable relative to each and every individual. But, again, I stay agnostic on what the justifiability standard requires.

The second premise moves from natural rights to natural duties. The natural right of basic justice and peace does not correlate with a natural duty to bring justifiable conventional practices into existence. (Who would have that duty?) Like the natural right of basic justice and peace, it is a bundle of Hohfeldian incidents—namely, the duties, no-rights, and liabilities that correlate with the rights that constitute the natural right of basic justice and peace. Thus the natural duty of basic justice and peace, the natural duty to respect other persons' natural right to participate in justifiable conventional practices that secure basic justice and peace, includes a duty not to exclude anyone from such practices, a no-right that others do not participate in such practices, and a liability to have one's claim rights and liberty rights changed when others acquire and exercise powers in line with them. This liability is key to explaining political authority. I have a few more things to say in defense of this natural duty of basic justice and peace in response to the third objection below.

The third premise says that the rules of justifiable conventional practices that secure peace and basic justice can confer the bundle of rights that constitutes political authority on the state (or, more precisely, on particular institutions that together constitute the state), including a content-independent power to impose duties on everyone on the territory and all citizens. For example, the rules of a conventional practice can determine that there is an institution such as a parliament that is equipped with certain rights, including the power to impose duties, and it can determine how one can become a member of parliament. As explained above, these rights are, first of all, legal rights; but if people have a natural right of basic justice and peace that allows them to acquire new claim rights, liberty rights, and powers in ways delineated by conventional practices that secure peace justice and peace, then these rights are morally validated: people who acquire them also acquire the corresponding moral rights.[28] Together, the rules of conventional practices can thus confer the moral rights that constitute political authority on institutions that together constitute a state.

In contrast to other natural duty accounts, the practice account need not claim that states are *necessary* to provide basic justice and peace. As explained above, accounts based on a natural duty of justice have to claim that institutions acquire political authority only when compliance is necessary to discharge one's natural duty of justice because otherwise justice-promoting institutions like, for example, Amnesty International would implausibly acquire political authority as well. The practice account can avoid ascribing political authority to such institutions because first, the rules that govern these institutions in fact do not confer political authority on them, and, second, they arguably would not be justifiable if they did. Accounts based on a natural duty of rescue have to claim that states are necessary to provide peace because otherwise the duty of rescue does not apply. The natural duty of basic justice and peace, on the other hand, applies to *all* practices that are justifiable and secure and maintain basic justice and peace, no matter if there are alternative practices that could do the job as well.

In other words, then, the practice account needs to claim only that justifiable conventional schemes can confer authority on states because this is *one way* to provide basic justice and peace. When the (justifiable) conventional practice that provides basic justice and peace for you involves state institutions, then you have a natural right to participate in *that* practice; when the (justifiable) conventional practice that provides basic justice and peace for you does not involve state institutions, then you have a natural right to participate in *that* practice. That the practice account can remain agnostic regarding the necessity of the state is an advantage, I think, in the absence of empirical evidence about the viability of ordered anarchy and given the maybe optimistic but reasonable

28 On the fair play account of political authority sketched above, the principle of fair play explains why the legal powers (and other legal rights) assigned by legal rules are morally validated. As Tosi puts it, "when people become subject to fair play schemes . . . , the rules of the scheme specify the terms of their cooperation. . . . In this way the principle of fair play turns the rules of the scheme (and the rights they specify) into moral requirements" ("The Possibility of a Fair Play Account of Legitimacy," 95). On the account proposed here, it is the natural duty and natural right of basic justice and peace that explain why the legal powers (and other legal rights) accorded by legal rules are morally validated.

arguments of anarchists.[29] Note that the claim that the state is *necessary* to provide basic justice and peace would be false even if states were much better than stateless societies in maintaining peace and justice, as long as the latter are *able* to maintain basic justice and peace.

The fourth step draws the conclusion that persons can acquire the rights associated with political authority, e.g., the rights of members of parliament, when these rights are conferred by the rules of a justifiable practice. The fifth step draws the mirroring conclusion that all citizens and everyone on a state's territory have a liability to have duties imposed by state institutions when the rules of a justifiable practice confer these rights on state institutions.

This, in a nutshell, is the practice account of political authority. On the natural duty accounts discussed above, the state has the authority to impose duties on persons because this allows persons to discharge their natural duties. On the practice account, in contrast, the state has the authority to impose duties on persons because this authority is conferred on the state by legal rules that persons have natural duties to respect.

5. Objections

The previous section introduces the practice account of political authority, which, unlike other natural duty accounts, offers a working mechanism that explains how and why states can realistically acquire political authority. I now discuss seven objections to the practice account. Replying to them gives me the chance to further clarify, defend, and elaborate the practice account.

5.1 The Practice Account Disregards the Distinction Between Justifiability and Legitimacy

The practice account disregards A. John Simmons's distinction between justifiability and legitimacy, according to a first objection: it is one thing to assess the moral qualities of an organization (i.e., its justifiability); it is another to assess which rights it has over me and others (i.e., its legitimacy).[30] This is of course a worry that is not peculiar to the practice account; it holds for all natural duty accounts of political authority. But it deserves a (brief) reply.

29 E.g., Friedman, *The Machinery of Freedom*; and Huemer, *The Problem of Political Authority*, part 2. Kantian functionalist accounts try to establish *a priori* that the state is necessary to realize justice. See Stilz, *Liberal Loyalty*, ch. 2; Ripstein, *Force and Freedom*, ch. 6; and Layman, "Rights, Respect, and Equality," 79–91. For recent criticism, see Christmas, "Against Kantian Statism."

30 Simmons, "Justification and Legitimacy" and "The Duty to Obey and Our Natural Moral Duties," 148–49.

Simmons is right that in general, one should distinguish between the moral virtues of an organization and the rights it has. But this is different for institutions that secure basic justice and peace, if the practice account is correct. The practice account indeed claims that if a state is justifiable—if it is part of a justifiable conventional practice that secures basic justice and peace—then the state also has the rights that constitute political authority.[31] And the natural duty of basic justice and peace is to explain why this is so. Of course, one may doubt that people actually have this natural duty of basic justice and peace. But that points to a different set of objections, which I get to next.

5.2. The Natural Duty of Basic Justice and Peace Is Implausibly Costly

Natural duties that require more than noninterference typically apply only when the costs of fulfilling them are reasonable.[32] Duties of rescue, for example, do not arise in scenarios where one would have to risk one's own life in order to have a chance to save someone else. Because this is so, it has been argued that one cannot ground a duty to obey the law in a natural duty of rescue because obedience to the law would be an unreasonable cost. Wellman replies that while the costs of obedience can be considerable, the benefits states provide are considerable too, such that the *net* costs of obedience are not unreasonable.[33] Against Wellman, Simmons and George Klosko point out that the benefits of the state are not causally linked to individual compliance and, for that reason, cannot be deducted from the costs.[34]

Applied to the practice account, the objection is that the natural duty of basic justice and peace needs a cost qualifier too if it is to be credible as a natural duty that involves more than noninterference. It should not make us respect other people's natural right of basic justice and peace no matter what, but only when this respect does not come with unreasonable costs. As such, it cannot contain an unqualified liability to have duties imposed when people acquire

31 From a different angle, one could suspect that there is a problematic circularity involved in my account: Can we even assess whether a state is justified without *presupposing* that it has political authority? I agree that states without political authority are deeply morally troublesome (Wendt, "Justice and Political Authority in Left-Libertarianism"). But I do not think this means that there is a circularity involved in my argument. Since on my account, states that are justified have political authority, we should not make political authority a *criterion* of justifiability.

32 Klosko, *Political Obligations*, 77, and "Fair Play, Reciprocity, and Natural Duties of Justice," 340–41.

33 Wellman, "Samaritanism and the Duty to Obey the Law."

34 Simmons, "The Duty to Obey and Our Natural Moral Duties," 182; and Klosko, "Fair Play, Reciprocity, and Natural Duties of Justice," 347. See also Klosko, *Political Obligations*, 94–95.

and exercise powers in line with the rules of justifiable practices that secure and maintain basic justice and peace. At best, it may contain a liability to have duties imposed *except when* this is an unreasonable cost for a duty bearer.

But in reply, there are actually *two* cost qualifiers implicit in the natural duty of basic justice and peace. The first is that it applies only to *justifiable* practices. For example, if more-than-minimal or more-than-small states are taken to impose unreasonable costs, then only minimal or small states have political authority, because only they will be justifiable, and so the natural duty of basic justice and peace does not apply to more-than-minimal or more-than-small states. As indicated above, I here stay agnostic on the details of the justifiability standard.

The second implicit cost qualifier is that the content-independent power to impose duties is not without moral limits. When a law is deeply unjust, it transgresses these moral limits, even if the state is endowed with political authority, such that no one incurs a duty to comply with it. For what it is worth, I think that the moral limits of the power to impose duties cannot be so narrow that they rule out even *mildly* unjust law.[35] But not much depends on this for our context. If you think that it is an unreasonable cost to incur a duty to comply with mildly unjust law, you still do not have to reject the natural duty of basic justice and peace—you can simply advocate a very narrow conception of the moral limits of the content-independent power to impose duties. This leaves the core idea of the practice account intact.

The remaining liability to have duties imposed is not, it seems to me, an unreasonable cost. It is the bare mirror of the natural right of basic justice and peace. If the natural duty of basic justice and peace did not contain the liability to have duties imposed, if the rules of a justifiable practice allow for that, then the natural right of basic justice and peace could also not contain the moral power to acquire new powers in line with the rules of such practices.

5.3. *The Natural Duty of Basic Justice and Peace Is a Philosophical Fantasy*

That the natural duty of basic justice and peace is not unreasonably costly may not be enough to convince you of its existence. Unlike Wellman's natural duty of rescue and Renzo's duty not to harm others, the natural duty of basic justice and peace that I rely on here is not part of commonsense morality.[36] What I

35 That persons have a duty to comply with mildly unjust law does not mean that they have to comply *all things considered*, of course. All duties are *pro tanto* duties, and the duty to comply with a mildly unjust law can of course be outweighed by the law's injustice. For recent discussion, see Brennan, *When All Else Fails*; Viehoff, "Legitimate Injustice and Acting for Others"; and Wellman, "The Space Between Justice and Legitimacy."

36 Simmons argues that Wellman's duty of rescue is actually a "curious hybrid" between a duty of rescue and a duty of charity since the state of nature is not an emergency in the

have to offer in support of the natural duty of basic justice and peace are the following thoughts.

First of all, if the argument of this article is sound, then we have to claim that people have a natural duty of basic justice and peace *if* we are to explain political authority. The natural duty of basic justice and peace is a *theoretical postulate* in that sense. And the case for postulating its existence is even stronger if one thinks that states need political authority to count as legitimate (as I do). If states need political authority to be legitimate, then one can read this article as spelling out what natural duty we have to postulate if at least some states are to count as legitimate. The alternative is to embrace political anarchism.

For the purposes of this article, this should be good enough. But I do have some additional thoughts to support the natural duty of basic justice and peace. Recall that the Hohfeldian duties, no-rights, and liabilities in the natural duty of basic justice and peace correlate with the claim rights, liberty rights, and powers in the natural right of basic justice and peace. One way to defend the existence of natural rights is by appealing to important interests that all humans share. Given the importance of everyone's interest in living in a society that provides basic justice and peace, I hope and think that the natural right of basic justice and peace and the corresponding natural duty of basic justice and peace have at least some initial plausibility.[37]

Perhaps one could argue that while people may have an interest in living a society that provides basic justice and peace, they do not have an interest in *participating* in such a practice. And it is indeed the participation that is crucial for my argument to take off. But of course a practice can be sustained only if people participate in it. What does this show? Well, first of all, if people have an interest in living in a society that provides basic justice and peace, they must also be taken to have an interest that at least *some* people participate and thereby sustain such a practice. Second, if at least some people are to rightfully participate in such a practice, at least some people must be conceived as endowed with a natural right to participate in such a practice. Finally, if we assume that people have the same natural rights, then if some must be conceived as endowed with

same way as a drowning child in a pond is ("The Duty to Obey and Our Natural Moral Duties," 182–84).

37 As far as I can see, this interest is not what Owens describes as a normative interest, an interest in "shaping the normative landscape" (*Shaping the Normative Landscape*). (See also Raz, *The Morality of Freedom*, 173.) It is simply an interest to live in a society that secures basic justice and peace; that some offices come with powers to shape the normative landscape is instrumental in securing basic justice and peace.

a natural right to participate in such a practice, everyone must be conceived as endowed with a natural right to participate in such a practice.[38]

5.4. States Are Not Justifiable If They Are Not Necessary for Basic Justice and Peace

A fourth objection raises doubts whether states can count as justifiable if they are not necessary to provide basic justice and peace, as the practice account concedes. States, after all, seem to violate rights when they raise funds via taxation and when they maintain a monopoly on the use of force. In a stateless society, rights may get violated too, to be sure. But this is a contingent feature of stateless societies. States, on the other hand, necessarily violate rights: it is a defining characteristic of states that they coercively uphold a monopoly on the use of force and coercively fund their operations. Since the practice account does not rely on the assumption that states are necessary to provide basic justice and peace, only stateless conventional practices that secure basic justice and peace should count as justifiable.

In response, we can reject that justifiable states *impermissibly* infringe on rights when they use taxation and maintain a monopoly on the use of force. Eric Mack proposes the following principle:

> When working out the detailed specification of persons' rights, one is to avoid specifications that systematically preclude individuals from exercising their rights or from conducting their lives in ways that a specification of their rights is supposed to protect.[39]

If persons' rights cannot be adequately protected without the state, they have to be attenuated such that they allow taxation and the state's monopoly on the use of force.[40] A specification of persons' rights can attenuate them in basically

38 My account has some similarity to the account of political obligation recently proposed by Valentini, who argues that we have a content-independent duty to obey the law because we have a moral duty to respect people's commitment to the rule of law when this commitment is authentic, morally permissible, and respecting it is not too costly (*Morality and Socially Constructed Norms*, 150, 167–70). She takes this to be a special instance of a quite general moral duty to respect socially constructed norms when they are authentically endorsed and morally permissible and when respecting them is not too costly (97–101). My account also highlights the role of conventions, and it also postulates an antecedent ("natural") moral duty. But otherwise, it is quite different. What underlies my account is people's fundamental interest in basic justice and peace, not respect for people's commitments; and it is not a Hohfeldian duty but a Hohfeldian liability (one part of the bundle of Hohfeldian incidents that I call the natural duty of basic justice and peace) that is to do the explanatory work.

39 Mack, "Nozickian Arguments for the More-than-Minimal State," 112.

40 Mack, "Nozickian Arguments for the More-than-Minimal State," 113–14.

two ways: it can say that (some) rights are sometimes permissibly infringed, and it can say that (some) rights apply only conditionally, such that they are not even infringed when they do not apply. When it comes to taxation, Mack argues that rights are permissibly infringed; in other contexts, he specifies rights conditionally.[41]

Now, Mack here involves a claim about the necessity of the state, which the practice account tries to avoid. The practice account claims that states can have political authority even if a stateless society could provide basic justice and peace as well. I do not think this undermines the basic idea behind Mack's principle, though: when *justifiable* states use taxation and maintain a monopoly on the use of force, they permissibly infringe upon rights. This does not mean that all taxation is fine. If, for example, only small states should count as justifiable—maybe because more extensive states are not in line with the point of the practice of private property—then only taxation within small states does not classify as an impermissible rights violation.[42]

An objector may insist that states cannot be justifiable if they infringe upon rights and are not even necessary to provide basic justice and peace. In reply: this is not so. States (or rather, certain kinds of states) can still be justifiable for several reasons: we lack evidence whether (and how well) nonstate institutions could provide basic justice and peace; it seems reasonable to suspect that states probably tend to be better at providing basic justice and peace than nonstate institutions; and the transition to nonstate institutions currently seems infeasible and would come with considerable risks.

5.5. *The Practice Account Is Incompatible with a Commitment to Moral Equality*

That we are moral equals means, at least, that we have equal rights. When some persons have authority over others, this constitutes an inequality of rights. Voluntary transfers of rights could justify this inequality, but the practice account does not rely on voluntary transfers. It allows some to unilaterally acquire and exercise authority over others. How is this supposed to be compatible with a commitment to moral equality understood as an equality of rights?

To develop a response, it is instructive to compare the unilateral duty imposition that the practice account allows to the unilateral duty imposition that comes with the acquisition of private property. Bas van der Vossen argues that the power to impose duties that is involved in the acquisition of private property is an instance of duty *alteration* rather than duty *creation*, which is why it

41 For discussion, see Wendt, "Political Authority and the Minimal State," 101–7.

42 For discussion, see Wendt, "Taxation and the Moral Authority of Conventions."

is compatible with everyone's moral equality.[43] Everyone has a natural duty to respect everyone else's natural right to engage in proper acts of appropriation, and a particular appropriation merely triggers that natural duty.

Similarly, on the practice account, everyone has a natural duty to respect everyone else's natural right of basic justice and peace, and a particular acquisition or exercise of powers associated with a political office merely triggers that natural duty. In that sense, the state's power to impose duties is a power to impose *new* duties only in a limited sense, if political authority is grounded in natural duties.[44] Of course, the state's power to impose duties is content independent (within moral limits), while the power to impose duties that is associated with the acquisition of private property is not. That is an important difference, and it arguably makes the former more worrisome than the latter. But it does not change the more fundamental point that the exercise of both powers involves a duty alteration against the background of an equality of moral rights and duties.

There is one complication: the natural right to participate in justifiable conventional practices gives you a moral right to acquire and exercise powers to impose duties in line with the rules of justifiable conventional practices *only if* these rules actually give everyone the chance to acquire such powers. If the rules of a practice allow only white men to get into political offices, for example, there is obviously no moral equality. Arguably, all *justifiable* practices today come with political positions that are open to (at least) all law-abiding adult citizens.

5.6. The Practice Account Does Not Confer Proper Authority

On the flip side, that political authority arises against a background of moral equality also points at a disadvantage of natural duty accounts in general and the practice account in particular: William Edmundson suggests that the doctor in the accident example has mere "leadership," which comes with considerably less error tolerance than proper authority.[45] The same holds for the state when its power to impose duties is tied to people's natural duties.

Similarly, Matthias Brinkmann distinguishes two models of legitimacy. According to the authority model, legitimate entities can *directly* impose duties on persons. According to the transmission model, legitimate entities can merely *indirectly* impose duties on persons: they "change the empirical

43 Van der Vossen, "Imposing Duties and Original Appropriation," 72–76.

44 Cf. Wendt, "Political Authority and the Minimal State," 117–21, and *Authority*, 73–75.

45 Edmundson, "Consent and Its Cousins," 344–46. See also Edmundson, "Political Authority, Moral Powers and the Intrinsic Value of Obedience," 188–90.

situation in which we find ourselves—for example, by establishing a salient signal of coordination—and this in turn triggers changes in our duties and rights."[46] Like Edmundson's "leadership," the power to indirectly impose duties is defeasible and closely tied to performance, and for that reason, it does not amount to a proper power to impose duties.[47] Natural duty accounts allow states to change the world such that people's natural duties are triggered in a way that gives rise to a liability to have duties imposed by the state, and so they fit the transmission model, as far as I can see.

I agree with Edmundson and Brinkmann that on the practice account, political authority is tied to performance and that it is a power to indirectly impose duties. But I deny that this means that it is not a proper power to impose duties or is less than proper political authority. What matters is that it is a content-independent power. And the practice account, in contrast to other natural duty accounts, offers a specific mechanism that explains how and why the state can acquire a content-independent power to impose duties (within moral limits)—namely, by having it conferred by the rules of justifiable practices.

5.7. The Practice Account Does Not Meet the Particularity Requirement

The most notorious problem for natural duty accounts of political authority (and political obligation) is to satisfy the so-called particularity requirement that was introduced into the debate by Simmons.[48] It naturally faces the practice account as well. Natural duties are duties we have in virtue of our humanity, not as Americans or Germans or Liberians. So how can natural duties explain that we are bound to comply with the laws of particular states but not others, and that states have political authority over particular persons and particular territories but not others? In the example of the accident discussed above, the doctor has authority over those who happen to be at the site and are able to help. But no comparable criterion seems available when it comes to political authority. This is, at core, the problem of satisfying the particularity requirement.

The most straightforward and pragmatic answer that proponents of natural duty (and functionalist) accounts can give is: first, states need discrete territorial jurisdiction to be able to do what they are supposed to do—i.e., to maintain justice and peace; second, while the current territorial divisions are morally arbitrary, they are salient; and third, because of their salience, we should treat current territorial divisions as authoritative when it comes to particularized

46 Brinkmann, *An Instrumentalist Theory of Political Legitimacy*, 46.

47 Brinkmann, *An Instrumentalist Theory of Political Legitimacy*, 47–49.

48 Simmons, *Moral Principles and Political Obligations*, 147–56, and "The Duty to Obey and Our Natural Moral Duties," 162–79.

natural duties.[49] One can then say with Daniel Layman that "the power of a government is legitimate over all of the people for whom it performs this service [of rectifying the defects of the state of nature]."[50]

But this pragmatic answer takes us only so far. First, while it may explain why states have authority over particular territories, it does not explain why states have authority over expatriate citizens.[51] Layman points out that states provide services to expatriate citizens.[52] But in contrast to discrete territorial jurisdiction, this certainly is not a pragmatic necessity for the ability to maintain justice and peace.[53] Second, the pragmatic answer is hard to reconcile with the impact of historical criteria: incidents of wrongful annexation, for example, seem to affect the authority that a state has over the annexed territory, no matter if the new territorial boundaries are now salient or not.[54]

In any case, the practice account has the resources to deal with the particularity problem in a different way.[55] The pragmatic answer is, so to say, built into the natural duty of basic justice and peace that is at the heart of the practice account: since everyone has a natural duty to respect everyone else's natural right to participate in justifiable conventional practices that secure basic justice and peace, these conventional practices can particularize everyone's liabilities.[56] When conventional practices are legal systems embedded in international law and bilateral treaties, they determine states' territories as well as the extent of their authority over citizens abroad (e.g., when it comes to taxation).

49 Cf. Waldron, "Special Ties and Natural Duties," 22–27; and Wellman, "Samaritanism and the Duty to Obey the Law," 37–46. Sometimes this argument is underpinned by a Kantian story about how persons are greater threats to those who are near them in the state of nature (cf. Waldron, "Special Ties and Natural Duties," 14–15); but this, it seems to me, does not contribute much to the justification of present boundaries (cf. Simmons, "The Duty to Obey and Our Natural Moral Duties," 170–76, and *Boundaries of Authority*, 68–73; and Klosko, "Fair Play, Reciprocity, and Natural Duties of Justice," 348).

50 Layman, "Rights, Respect, and Equality," 95.

51 Mason, "Special Obligations to Compatriots," 437.

52 Layman, "Rights, Respect, and Equality," 95–96.

53 One can of course bite the bullet and deny that states have authority over citizens who are not on their territory. See Wellman, "Samaritanism and the Duty to Obey the Law," 46–52. See also Zhu, "Content-Independence and Natural-Duty Theories of Political Obligation," 62.

54 Simmons, "The Duty to Obey and Our Natural Moral Duties," 173–74, and *Boundaries of Authority*, 51–52, 71, 75–80, 85.

55 On how to incorporate historical concerns into natural duty (and functionalist) accounts of political authority, see Stilz, "Territorial Boundaries and History"; and Motchoulski, "Justice, Reciprocity, and the Boundaries of State Authority," 62–69.

56 Similarly, Nieswandt argues that the addressee of a duty is determined by the rules of the relevant practice ("What Is Conventionalism About Moral Rights and Duties?" 22).

This does not mean that conventional practices have carte blanche to determine these matters in any way, of course. The natural right of basic justice and peace and natural duty of basic justice and peace appeal to *justifiable* practices. All kinds of moral concerns can affect the justifiability of a practice. This includes historical concerns: instances of annexation, for example, can undermine the justifiability of a state and thereby also undermine its political authority.

There is another point to consider. The practice account can be developed in a geographically holistic or a nonholistic version. In a holistic version, a practice as a whole is to be evaluated as either justifiable or not, and if it is justifiable, then all institutions that are part of the practice have the rights that the rules of the practice confer on them. The problem with that approach is that practices can be geographically individuated in different ways and in fact often seem to be overlapping. For example, state law is embedded in federal law in the United States, and in some contexts, it makes sense to conceive of a state like Virginia as a practice of its own, while in other contexts it makes more sense to conceive of it as part of the practice that is the United States. For that reason, I think a nonholistic approach is more sensible. On that approach, conventional practices that are part of a larger practice can count as justifiable, even if the larger practice is not justifiable, and vice versa. Accordingly, then, the practice account may be able to regard a state as justifiable and endowed with authority with regard to its original territory but not with regard to a newly annexed territory.[57]

Virginia Tech
fabianwendt@vt.edu

References

Adams, N. P. "Institutional Legitimacy." *Journal of Political Philosophy* 26, no. 1 (2018): 84–102.

Applbaum, Arthur I. "Legitimacy Without the Duty to Obey." *Philosophy and Public Affairs* 38, no. 3 (2010): 215–39.

57 I would like to thank two anonymous reviewers for *JESP* for their excellent comments; audiences at the American Political Science Association Annual Meeting in Los Angeles in 2023, the philosophy brown bag series at Virginia Tech in 2023, Laura Valentini's Colloquium in Social and Political Philosophy in Munich in 2023, the 7th Workshop in Political Philosophy in Graz in 2025, and the Tokyo Forum for Analytic Philosophy in 2025 for helpful discussions; Travis Quigley for his comments in Los Angeles; and Michael Moehler for comments on an early draft.

Brennan, Jason. *When All Else Fails: The Ethics of Resistance to State Injustice.* Princeton University Press, 2018.
Brinkmann, Matthias. *An Instrumentalist Theory of Political Legitimacy*. Oxford University Press, 2024.
Buchanan, Allen. "Political Legitimacy and Democracy." *Ethics* 112, no. 4 (2002): 689–719.
Christiano, Thomas. *The Constitution of Equality: Democratic Authority and Its Limits*. Oxford University Press, 2008.
Christmas, Billy. "Against Kantian Statism." *Journal of Politics* 83, no. 4 (2021): 1721–33.
Copp, David. "The Idea of a Legitimate State." *Philosophy and Public Affairs* 28, no. 1 (1999): 3–45.
Edmundson, William. "Consent and Its Cousins." *Ethics* 121, no. 2 (2011): 335–53.
———. "Political Authority, Moral Powers and the Intrinsic Value of Obedience." *Oxford Journal of Legal Studies* 30, no. 1 (2010): 179–91.
Enoch, David. "Authority and Reason-Giving." *Philosophy and Phenomenological Research* 89, no. 2 (2014): 296–332.
Estlund, David. *Democratic Authority*. Princeton University Press, 2008.
Friedman, David. *The Machinery of Freedom: Guide to a Radical Capitalism*. 2nd ed. Open Court, 1989.
Hart, H. L. A. "Are There Any Natural Rights?" *Philosophical Review* 64, no. 2 (1955): 175–91.
Hohfeld, Wesley N. *Fundamental Legal Conceptions As Applied in Judicial Reasoning*. Edited by David Campbell and Philip Thomas. Ashgate, 2001.
Huemer, Michael. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. Palgrave Macmillan, 2013.
Kavka, Gregory S. *Hobbesian Moral and Political Theory*. Princeton University Press, 1986.
Klosko, George. "Fair Play, Reciprocity, and Natural Duties of Justice." *Ratio Juris* 33, no. 4 (2020): 335–50.
———. *Political Obligations*. Oxford University Press, 2005.
———. *The Principle of Fairness and Political Obligation*. Rowman and Littlefield, 1992.
Koltonski, Daniel. "Normative Consent and Authority." *Journal of Moral Philosophy* 10, no. 3 (2013): 255–75.
Kramer, Matthew H. "Rights Without Trimmings." In *A Debate over Rights: Philosophical Enquiries*, by Matthew H. Kramer, N. E. Simmonds, and Hillel Steiner. Oxford University Press, 1998.
Layman, Daniel. "Rights, Respect, and Equality: The Basis of Authority." In *Is Political Authority an Illusion? A Debate*, by Michael Huemer and Daniel

Layman. Routledge, 2022.
Lefkowitz, David. "Simmons' Critique of Natural Duty Approaches to the Duty to Obey the Law." *APA Newsletter on Philosophy and Law* 7, no. 1 (2007): 9–14.
Mack, Eric. "Nozickian Arguments for the More-than-Minimal State." In *Cambridge Companion to Anarchy, State, and Utopia*, edited by Ralf M. Bader and John Meadowcroft. Cambridge University Press, 2011.
Mason, Andrew. "Special Obligations to Compatriots." *Ethics* 107, no. 3 (1997): 427–47.
Melenovsky, C. M. "The Reasons to Follow Conventional Practices." *Australasian Journal of Philosophy* 102, no. 3 (2024): 710–25.
Motchoulski, Alexander. "Justice, Reciprocity, and the Boundaries of State Authority." *Journal of Political Philosophy* 30, no. 1 (2022): 48–69.
Nieswandt, Katharina. "What Is Conventionalism About Moral Rights and Duties?" *Australasian Journal of Philosophy* 97, no. 1 (2019): 15–28.
Owens, David. *Shaping the Normative Landscape*. Oxford University Press, 2012.
Perry, Stephen. "Political Authority and Political Obligation." *Oxford Studies in Philosophy of Law* 2 (2013): 1–74.
Quong, Jonathan. *Liberalism Without Perfection*. Oxford University Press, 2011.
Rawls, John. *A Theory of Justice*. Harvard University Press, 1971.
Raz, Joseph. *The Morality of Freedom*. Clarendon Press, 1986.
Renzo, Massimo. "State Legitimacy and Self-Defence." *Law and Philosophy* 30, no. 5 (2011): 575–601.
Ripstein, Arthur. *Force and Freedom: Kant's Legal and Political Philosophy*. Harvard University Press, 2009.
Simmons, A. John. *Boundaries of Authority*. Oxford University Press, 2016.
———. "The Duty to Obey and Our Natural Moral Duties." In *Is There a Duty to Obey the Law?* by Christopher H. Wellman and A. John Simmons. Cambridge University Press, 2005.
———. "Justification and Legitimacy." In *Justification and Legitimacy: Essays on Rights and Obligations*, by A. John Simmons. Cambridge University Press, 2001.
———. *Moral Principles and Political Obligations*. Princeton University Press, 1979.
Sreenivasan, Gopal. "Oh, but You Should Have: Estlund on Normative Consent." *Iyyun* 58 (2009): 62–72.
Stilz, Anna. *Liberal Loyalty: Freedom, Obligation, and the State*. Princeton University Press, 2009.
———. "Territorial Boundaries and History." *Politics, Philosophy, and Economics* 18, no. 4 (2019): 374–85.
Tosi, Justin. "The Possibility of a Fair Play Account of Legitimacy." *Ratio* 30,

no. 1 (2017): 88–99.
Valentini, Laura. *Morality and Socially Constructed Norms*. Oxford University Press, 2023.
van der Vossen, Bas. "Imposing Duties and Original Appropriation." *Journal of Political Philosophy* 23, no. 1 (2015): 64–85.
Viehoff, Daniel. "Legitimate Injustice and Acting for Others." *Philosophy and Public Affairs* 50, no. 3 (2022): 301–74.
Waldron, Jeremy. "Special Ties and Natural Duties." *Philosophy and Public Affairs* 22, no. 1 (1993): 3–30.
Wellman, Christopher H. "Liberalism, Samaritanism, and Political Legitimacy." *Philosophy and Public Affairs* 25, no. 3 (1996): 211–37.
———. "Samaritanism and the Duty to Obey the Law." In *Is There a Duty to Obey the Law?* by Christopher H. Wellman and A. John Simmons. Cambridge University Press, 2005.
———. "The Space Between Justice and Legitimacy." *Journal of Political Philosophy* 31, no. 1 (2023): 3–23.
———. "Toward a Liberal Theory of Political Obligation." *Ethics* 111, no. 4 (2001): 735–59.
Wenar, Leif. "The Nature of Rights." *Philosophy and Public Affairs* 33, no. 3 (2005): 223–53.
Wendt, Fabian. *Authority*. Polity Press, 2018.
———. "Justice and Political Authority in Left-Libertarianism." *Politics, Philosophy, and Economics* 14, no. 3 (2015): 316–39.
———. "Political Authority and the Minimal State." *Social Theory and Practice* 42, no. 1 (2016): 97–122.
———. "Taxation and the Moral Authority of Conventions." *Social Philosophy and Policy* 39, no. 1 (2023): 118–38.
Zhu, Jiafeng. "Content-Independence and Natural-Duty Theories of Political Obligation." *Philosophy and Social Criticism* 44, no. 1 (2018): 61–80.
———. "Farewell to Political Obligation: In Defense of a Permissive Conception of Legitimacy." *Pacific Philosophical Quarterly* 98, no. 3 (2017): 449–69.

Journal of Ethics and Social Philosophy
vol. 30, no. 6 · october 2025

https://doi.org/10.26556/jesp.v30i6.3246

SUPERSESSION-PROOF REPARATIONS

HARMS, WRONGS, AND HISTORICAL INJUSTICE

Felix Lambrecht

There is widespread intuition that historical injustices require some form of redress. Despite this intuition, redress for historical injustice encounters significant philosophical problems. In this article, I defend the possibility of redress from one particular philosophical problem: the *supersession thesis*. According to the supersession thesis, circumstances may have changed between the historical injustice and the present such that present demands of justice override or "supersede" demands of redress for the historical injustice.[1] The supersession thesis challenges the possibility of redress for historical injustice. However, I argue that the way the supersession thesis challenges redress helpfully informs the model of redress for historical injustice that we should adopt.

There are two models of how an injustice generates a claim to redress:

> *Wrongful Injury Claim (WIC)*: Agent *A* has a claim to redress for wrongful action ϕ iff *A* experiences an injury *I* that is the result of ϕ.[2]

> *Wrongful Action Claim (WAC)*: Agent *A* has a claim to redress for wrongful action ϕ iff *A* is wronged by ϕ.[3]

1 The supersession thesis was originally presented by Waldron in "Superseding Historic Injustice" and has been frequently discussed in the historical injustice literature. A representative sample of these discussions includes Nine, "Superseding Historic Injustice and Territorial Rights"; Sanderson, "Redressing the Right Wrong"; Meyer and Waligore, "Supersession and Compensation for Historical Injustice"; Song, "Superseding Structural Linguistic Injustice"; Montero, "Colonialism and Rights Supersession"; Christie, "The Supersession of Indigenous Understandings of Justice and Morals"; and Harrison, "Supersession, Reparations, and Restitution."

2 A representative sample of discussions that use WIC includes Sher, "Ancient Wrongs and Modern Rights" and "Transgenerational Compensation"; Boxill, "A Lockean Argument for Black Reparations"; Wenar, "Reparations for the Future"; and Harrison, "Supersession, Reparations, and Restitution."

3 A representative sample of discussions that use WAC includes Butt, "Nations, Overlapping Generations, and Historic Injustice"; Shiffrin, "Reparations for US Slavery and Justice

In this article, I argue that the WIC model is vulnerable to the supersession thesis, while the WAC model is not. Thus, if we want an account of redress for historical injustice that overcomes the supersession thesis, we must prefer the WAC model over the WIC model.

I begin in section 1 by explaining the supersession thesis and showing how it challenges the possibility of redress for historical injustice. Then, in section 2, I present the WIC model in greater detail. In section 3, I consider a recent argument by Caleb Harrison that offers the best version of the WIC model that is meant to overcome the supersession thesis. However, I argue in section 4 that even this best version of the WIC model overcomes the supersession thesis only at the expense of significant further objections. The WIC model overcomes the supersession thesis by relying on injuries that causally result from the wrongs of the historical injustice. I argue that this reliance on injuries makes the WIC model vulnerable to the *nonidentity problem* and the *causal problem*. Finally, in section 5, I demonstrate how the WAC model can harness the insights of the WIC model to overcome the supersession thesis while at the same time avoiding the nonidentity and causal problems.

A few clarificatory comments about the goals of this article before I begin. First, both WIC and WAC are models of what generates a claim to redress. Neither tells us *how much* the redress should be or what it ought to be made up of. The question of what generates a claim is separate from the question of how much redress there ought to be.[4] My focus here is the possibility of *any* redress for historical injustice. Second, my goal in this article is not to develop the WAC model. Many others have developed versions of this model.[5] My contribution here is to demonstrate how the WAC model overcomes the supersession thesis by showing how it can incorporate the insights from the best version of the WIC model while avoiding its shortcomings. Thus, I explain only the details of the WAC model that are necessary to illustrate how it can overcome the supersession thesis. Third, philosophers have provided many recent arguments meant to overcome the supersession thesis. The majority of these arguments, however, aim to challenge the idea that circumstances can change what justice requires

over Time"; Sanderson, "Redressing the Right Wrong"; Kumar, "Why Reparations?"; and Lambrecht, "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

4 For discussion of this second question, see Goodin, "Theories of Compensation"; Sher, "Ancient Wrongs and Modern Rights"; Simmons, "Historical Rights and Fair Shares"; and Lazar, "Corrective Justice and the Possibility of Rectification."

5 See, e.g., Kumar, "Why Reparations?"; Shiffrin, "Reparations for US Slavery and Justice over Time"; Sanderson, "Redressing the Right Wrong"; Simmons, "Historical Rights and Fair Shares"; and Lambrecht, "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

in the way the supersession thesis says or to demonstrate that the supersession thesis does not arise in particular cases.[6] My argument tries to overcome the supersession thesis *on its own terms*. I aim to vindicate the possibility of redress for historical injustice even if we accept the claim in the supersession thesis that changes in present circumstances change what justice demands. That is, my goal is to show that even if we accept the parameters of the supersession thesis for the sake of argument, claims to redress for historical injustices are possible. Fourth, there are many arguments that groups *qua* groups have claims to redress for historical injustice (for instance, the claims of Indigenous nations).[7] However, group claims face significant complications. And the majority of the arguments I discuss here concern claims of individuals. I thus restrict my discussion to claims that individuals may have. Finally, my topic in this article is *backward-looking redress*. Recently, literature on historical injustice has taken a "structural turn," emphasizing the importance of forward-looking measures to repair unjust distributions or structures that have resulted from historical injustice.[8] The structural turn is welcome, and I have defended it elsewhere.[9] Yet as I and others argue, an *optimal* account of redress for historical injustice must also include backward-looking redress.[10] The supersession thesis challenges backward-looking redress. And so part of vindicating an optimal account of redress requires demonstrating that backward-looking redress is not vulnerable to the

6 For instance, Meyer and Waligore distinguish between "partial" and "full" supersession to argue that in many cases of historical injustice, the circumstances have not changed such that demands of present justice supersede claims of redress ("Supersession and Compensation for Historical Injustice"). Montero challenges the range of cases in which the supersession thesis applies by suggesting that it applies only if agents do not experience the minimum requirements of justice ("Colonialism and Rights Supersession"). Nine and Christie argue that some kinds of wrongs (e.g., land expropriations) cannot be superseded (Nine, "Superseding Historic Injustice and Territorial Rights"; Christie, "The Supersession of Indigenous Understandings of Justice and Morals"). Song and Lu argue that the supersession thesis does not apply because historical injustices are structural and enduring, so circumstances have not sufficiently changed (Song, "Superseding Structural Linguistic Injustice?"; Lu, *Justice and Reconciliation in World Politics*, 200).

7 E.g., Butt, "Nations, Overlapping Generations, and Historic Injustice"; and Thompson, "Groups as Intergenerational Agents."

8 See, e.g., Nuti, *Injustice and the Reproduction of History*; Lu, *Justice and Reconciliation in World Politics*; and Butt, "What Structural Injustice Theory Leaves Out."

9 See Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Pluralism, Structural Injustice, and Reparations for Historical Injustice."

10 See Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Reparations and Structural Injustice." See also Song, "Denial of Japan's Military Sexual Slavery and Responsibility for Epistemic Amends"; Butt, "What Structural Injustice Theory Leaves Out"; and McKeown, "Backward-Looking Reparations and Structural Injustice."

supersession thesis. For the remainder of this article, when I refer to redress or claims to redress, I refer to backward-looking redress unless otherwise specified.

1. INTRODUCING SUPERSESSION

The supersession thesis states that when circumstances change, claims to redress some historical injustice no longer apply because present claims of justice "supersede" historical claims.[11] Jeremy Waldron's influential argument for the supersession thesis is as follows. Suppose that at t_1, one group of individuals (*A*'s) has rights to a piece of land. Another group (*B*'s) wrongfully expropriates this land. At t_1, this was an injustice. But now, many generations later at t_2, the *B*'s have made their lives on this land and depend on this land for survival. The *B*'s have acquired a right to this land.[12] At t_1, the *A*'s had claims to redress for the injustice (in the form of return of the land). But at t_2, the *B*'s have present claims to this land. Taking the land would constitute a serious wrong and violate the rights of the *B*'s to it in the present. Present demands of justice, in other words, make it such that demands for redress cannot be met. Thus, the claims of the *A*'s to redress for historical injustice have been "superseded" by the claims of the *B*'s to the land now.

The supersession thesis seems to challenge redress in many (if not most) cases of historical injustice. Typically, philosophers treat historical injustices as collections of unjust actions that occurred generations ago between individual wrongdoers and victims who are no longer alive.[13] While there is a sense in which any unjust action committed in the past is "historical," philosophers typically are concerned with a particular understanding of historical injustice. Historical injustices are philosophically challenging because we cannot apply a standard model of redress for them. The individuals involved no longer exist, and the circumstances have changed since the injustice. Wrongs that occur relatively recently are not philosophically challenging in this same way. So when philosophers talk of historical injustices, they typically understand them as *philosophically challenging* injustices committed generations ago between indi-

11 Waldron "Superseding Historic Injustice," 24.

12 Waldron "Superseding Historic Injustice," 18.

13 Many philosophers define historical injustice in this way, e.g., Butt, "Nations, Overlapping Generations, and Historic Injustice," 358; Thompson, "Historical Injustice and Reparation," 116; Waldron, "Superseding Historic Injustice, 6–8; and Wenar, "Reparations for the Future," 399–401. I have elsewhere offered a detailed explanation of this understanding of historical injustice ("On the Necessity of a Pluralist Theory of Reparations for Historical Injustice"). Thanks to an anonymous reviewer for encouraging me to expand upon this point here.

vidual wrongdoers and victims who are no longer alive. This understanding does not preclude *also* understanding the wrongs of historical injustice as enduring or ongoing.[14] The central point is that to be philosophically interesting, *some of the actions* of the historical injustice must occur in the distant past. The standard examples of such historical injustices are chattel slavery in the United States and the land expropriations of Indigenous peoples in North America, Australia, and New Zealand. I restrict my discussion here to such wrongs not because I think they are the only possible historical injustices but rather because they are the central examples that the literature uses, they helpfully illustrate the philosophical problem with claims to redress, and my main interlocuters in this article use them.

Understanding historical injustice in this way is also consistent with the supersession thesis: the supersession thesis seems to apply most plausibly to injustices occurring in the distant past. This is because the supersession thesis seems to require quite a lot of time to pass in order for the circumstances to change in the relevant way. Consider Waldron's case of land expropriation. For the *B*'s to have acquired rights to the land, Waldron argues that they must have spent considerable time on the land and made it central to their families' identity through generations.[15] In other words, for the land to belong the *B*'s such that taking it away would cause a problem of present injustice, a lot of time must have passed. So to fairly defend the possibility of redress from the challenge of the supersession thesis, we need to use the same kinds of cases that best serve the supersession thesis (i.e., injustices in the distant past). I therefore restrict my discussion here to such cases. For the remainder of this article, when I refer to historical injustices, I mean wrongful actions committed generations ago between individual wrongdoers and victims who are no longer alive.

The supersession thesis seems to challenge many claims to redress for historical injustice. For the purposes of this article and for the sake of argument, I accept the idea that changes to circumstances can create present demands of justice in the way the supersession thesis states (for instance, by changing property rights). And for the sake of argument, I accept that these present demands of justice may change what claims to redress might require. I now turn to a plausible argument that seems to vindicate claims to redress. While I ultimately show that this argument fails, it does so instructively and in a way that gives us resources for a successful defense against the supersession thesis.

14 Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice."

15 Waldron, "Superseding Historical Injustice," 23–24.

2. The Wrongful Injury Claim Model

Recall from above that there are two main models that explain how wrongs generate claims to redress. One of these is the wrongful injury claim (WIC) model. Recently, Caleb Harrison has offered a strong defense of the possibility of redress for historical injustice against the challenge of the supersession thesis using the WIC model.[16] Let us consider it in detail.

Recall, the WIC model is as follows:

> Agent A has a claim to redress for wrongful action ϕ iff A experiences an injury I that is the result of ϕ.

The WIC model is widely adopted.[17] It has a few important features to note. First, proponents of the WIC model typically define an injury as an agent's well-being departing from a baseline. Usually, this baseline is put in counterfactual terms such that an agent is injured when that agent is worse-off than she would have been had the injustice not occurred.[18] I discuss differences in the baseline in section 4.1.2 below. For now, the point is that we can follow the WIC model and define injury in terms of a drop in well-being below some (typically counterfactual) baseline.

Second, not every action that results in an injury is a wrong, and not every wrong necessarily results in injury. An injury occurs when an action drops an agent's well-being. A *wrong* or *wrongful action* (I use these terms interchangeably here) is an action that is *impermissible*. Different accounts vary in the reasons that make an action impermissible. An action may be impermissible when it violates a right, when it is contrary to normative expectations, when it expresses something about the victim, when it inflicts a particular kind of injury on a victim, or when it violates a duty owed to a victim for some other reason.[19] I remain neutral about what makes an action wrong. The crucial point is that injuries and wrongs can come apart: an action can inflict an injury and not be wrong (e.g., when I step on your foot in order to save a child's life) and can be

16 Harrison, "Supersession, Reparations, and Restitution."

17 E.g., Boxill, "A Lockean Argument for Black Reparations"; Sher, "Ancient Wrongs and Modern Rights" and "Transgenerational Compensation"; Wenar, "Reparations for the Future"; and Harrison, "Supersession, Reparations, and Restitution."

18 See Boxill, "A Lockean Argument for Black Reparations," 67; Sher, "Transgenerational Compensation," 181–83; and Harrison, "Supersession, Reparations, and Restitution, 7–8.

19 For rights views, see Ripstein, *Force and Freedom* and *Private Wrongs*; Weinrib, *The Idea of Private Law*; and Hurley and Weinberg, "Whose Problem Is Non-Identity?" For normative expectations, see Kumar, "Who Can Be Wronged?" and "Why Reparations?" For expressive views, see Anderson and Pildes, "Expressive Theories of Law." For reasons, see Gardner, "What Is Tort Law For?"

wrong without inflicting an injury (e.g., when I scroll through your phone when you leave the room, and you never find out). This allows for the possibility of "harmless wrongdoing."[20] A famous example is offered by Arthur Ripstein: a man trespasses into a house and sleeps in the owner's bed while the owner is away without the owner ever finding out.[21] Such cases of "harmless wrongdoing" involve actions that are intuitively wrong, but their wrongfulness cannot be (fully) explained in terms of the injuries they cause (since they cause none). There is something else that makes these actions wrong. The WIC model insists that an action ϕ generates a claim to redress if and only if ϕ is a wrong, *and* ϕ causally results in an injury.

Third, the WIC model is neutral about the agent who must provide redress. Most philosophers who adopt the WIC model argue that the wrongdoer needs to provide redress.[22] This is what makes it redress and not merely third-party compensation. For simplicity, I assume that the agent who is required to provide redress is the wrongdoer. But this does not make a difference to my argument.

3. Can the Wrongful Injury Claim Model Overcome the Supersession Thesis?

We can now understand how the WIC model can be used to overcome the supersession thesis. The best attempt to do so is offered by Harrison.[23] Harrison does not dispute that the supersession thesis might challenge *some form* of redress for historical injustice. Rather, Harrison argues that the supersession thesis does not challenge redress *entirely*. He distinguishes between two forms of redress: *restitution* and *reparation*. Restitution is the return of the precise thing lost in the injustice.[24] For instance, restitution would be the return of the expropriated land. Reparation, in contrast, consists in payments or measures

20 For a representative sample of discussions of harmless wrongdoing, see Feinberg, *Harmless Wrongdoing*; Ripstein, *Force and Freedom*; Weinrib, *The Idea of Private Law*; Hurley and Weinberg, "Whose Problem Is Non-Identity?"; Kumar "Who Can Be Wronged?" and "Why Reparations?"; and Slavny and Parr, "Harmless Discrimination."

21 Ripstein, *Force and Freedom*, 125.

22 See, e.g., Boxill, "A Lockean Argument for Black Reparations," 64–66; Harrison, "Supersession, Reparations, and Restitution," 9–10; and Boxill and Corlett, "Black Reparations."

23 Harrison, "Supersession, Reparations, and Restitution." Meyer and Waligore offer a similar argument ("Supersession and Compensation for Historical Injustice"). They focus more on an external critique of Waldron's argument, so here I focus on Harrison's argument. Importantly, Meyer and Waligore also use the WIC model (see, e.g., p. 21), and so their argument is vulnerable to the same problems that I raise for Harrison's.

24 Harrison "Supersession, Reparations, and Restitution," 8.

that a wrongdoer must provide to the victim in order to address the wrong. Reparations can include material payments or practices of symbolic repair.[25] An injustice can call for some combination of restitution and reparation. For instance, when I steal your bicycle and keep it for a month, I owe you restitution (the return of the bicycle) and reparations (e.g., payment for money you spent on public transit instead of riding your bicycle to work). Crucially, restitution and reparation can come apart: redress for a wrong might call for one but not the other, and we can exercise them independently. This might occur when restitution is no longer possible. Suppose I steal your bicycle and destroy it. Restitution is now impossible. But I still owe you reparations to address the injustice in some way, for instance, with money and an apology.

Harrison argues that while the supersession thesis might challenge claims to restitution, it does not challenge claims to reparation. And since restitution and reparation come apart, the supersession thesis does not entirely challenge the possibility of claims to redress. Recall the example of land expropriation. When *B* expropriates *A*'s land, and then *B* makes a life on it, *A*'s claim to restitution may be superseded by *B*'s claim to the land. That is, as the supersession thesis says, at t_2 (the present), *B* may indeed have acquired legitimate title and made a life on this land such that returning the land to *A* may cause a greater injustice to *B*. However, even if we cannot return the land and thus *A* may not have a claim to restitution, *A* may still have a claim to *reparation* for *B*'s wrong. While *B* does not have to give the land back, *B* may still have to pay *B* for use of the land, apologize for taking it, and reinstate institutions of self-governance that *B* had on the land.[26]

Harrison's argument gives us an important tool. By distinguishing between restitution and reparations as kinds of redress, we may be able to overcome the supersession thesis. While demands of present justice may supersede the possibility of some kinds of redress (e.g., restitution), they may not supersede *all* kinds of redress (e.g., reparation). However, in making this argument, Harrison explicitly uses the WIC model.[27] That is, on his argument, what generates claims to reparations and redress are *injuries*. In using the WIC model to overcome the supersession thesis, Harrison's solution opens itself up to two further problems that ultimately challenge the possibility of redress entirely.

25 Harrison "Supersession, Reparations, and Restitution," 7–9.

26 Harrison "Supersession, Reparations, and Restitution," 12–13.

27 Harrison, "Supersession, Reparations, and Restitution," 7–10.

4. Problems for the Wrongful Injury Claim Model

The WIC model can overcome the supersession thesis because it makes injury a necessary condition for a claim to redress: if a party is injured, the agent is owed redress at least in the form of reparation to address this injury even if restitution is impossible. However, by making an injury a necessary condition for redress (and reparation), this solution opens itself up to two problems: the *nonidentity objection* and the *causal problem*.

4.1. The Nonidentity Objection

The nonidentity problem is a general philosophical puzzle developed by Derek Parfit that runs as follows.[28] When an event is a necessary condition for an agent coming to exist, that agent does not have a moral claim against this event because if the event had not happened, the agent would not have come to exist. Assuming the agent lives a life worth living (which I assume for the rest of this article), the agent cannot be said to be made worse-off by the event.

The nonidentity problem challenges claims to redress for historical injustice generated by the WIC model, which says that a necessary condition for any agent to have a claim to redress for a wrong is that that agent experiences an injury resulting from that wrong.[29] But according to the nonidentity problem, no agent in the present is made worse-off by the historical injustice because historical injustices are necessary conditions for every present agent having come to exist. So no present agent is injured by the injustice, and thus, no present agent is owed redress for historical injustices. Call this the nonidentity objection to redress for historical injustices. Put precisely:

- P1. An agent has a claim to redress for a wrong iff the agent is injured by that wrong. (WIC Model)
- P2. If an event is a necessary condition for an agent coming to exist, then this agent cannot be said to be injured by this event. (Nonidentity Problem)
- P3. Historical injustices are necessary conditions for all present agents having come to exist. (Empirical Observation)
- P4. No present agent is injured by any historical injustice. (From P2 and P3)
- C. No present agent is owed redress justice for historical injustices. (From P1 and P4)

28 See Parfit, *Reasons and Persons*.

29 The nonidentity problem was first raised in the context of historical injustice by Morris, "Existential Limits to the Rectification of Past Wrongs."

I will now defend each of these premises.

4.1.1. Defense of P1

P1 simply assumes the WIC model explained above.

4.1.2. Defense of P2

P2 says that an action does not injure an agent if that agent would not have existed without this action. P2 seems to assume a *counterfactual account* of injury, and most proponents of the WIC model assume this theory of injury.[30] The counterfactual account of injury says that an action makes an agent worse-off (injures an agent) only if that agent is worse-off than that agent would have been had the action not occurred.[31] P2 seems to assume this account: if an event is a necessary condition for an agent coming to exist, this agent is not made worse-off than the agent would have been by the action, since the agent would not have come to exist without this action.

It might be thought, then, that one way to object to P2 is to reject the counterfactual account of injury. However, we cannot reject P2 simply by rejecting the counterfactual account and replacing it with an alternative plausible account of injury. To reject P2 by providing an alternative account of injury, an objector must also demonstrate that this alternative account of injury does better than the counterfactual account and can generate claims to redress for the historical injustice. No account of injury can do this. My argument in this section is not about whether any of the accounts of injury I consider are the correct general account of injury. Rather, my argument is that alternative accounts of injury either cannot be incorporated into a plausible account of the WIC model or cannot generate claims to redress in cases of historical injustice in a way that fares any better than the counterfactual account implied by P2.

The counterfactual account is a *comparative* account of injury. It says that an agent is injured by an action when the agent is made worse-off by this action *compared to a baseline* where this action would not have occurred. The non-identity problem arises because of this comparative baseline. If the action had not occurred, then the agent would not exist. The agent cannot be said to be worse-off compared to the baseline in which the action had not occurred. There is no comparative baseline where the victim would have both existed and been

30 E.g., Boxill, "A Lockean Argument for Black Reparations," 67; Sher "Transgenerational Compensation," 181–83; Wenar, "Reparations for the Future," 397–99; and Harrison, "Supersession, Reparations, and Restitution," 7–8.

31 See Boonin, *The Non-Identity Problem and the Ethics of Future People*; Purves, "Harming as Making Worse Off"; Johansson and Risberg, "A Simple Analysis of Harm"; and Carlson et al., "Plural Harm."

better-off. Any other comparative account of injury has this problem.[32] Take, for instance, the *temporal account* of injury. On the temporal account, an agent is injured by an action if and only if the action makes the agent worse-off than the agent was before the action was committed. However, even though this account does not need to appeal to a counterfactual baseline, it still appeals to the temporal baseline (the well-being at the time right before the action was committed). As with the counterfactual account, the comparative baseline is not possible. Before the action was committed, the agent did not exist. And so we cannot say that this agent is worse-off than at any time before the action was committed.

One might suggest a *moralized* comparative account of injury. A moralized account says that an action injures an agent when the action makes the agent worse-off than the agent *ought* to be or ought to have been. This appears to be a comparative account of injury that does not rely on a baseline that is vulnerable to the nonidentity problem. Even if an agent is not better-off than the agent would have been had the action not been committed, it seems as though we can say that the agent ought to have been *even better-off*. But this moralized "worse-off than one ought to have been" is ambiguous between two understandings. On the first understanding, an agent is made worse-off than the agent ought to have been, where "ought" is understood *in comparative terms to a different action that could have been committed*. On the second understanding, an agent is made worse-off than the agent ought to have been, where "ought" is understood in terms of *independent moral reasons*. The first comparative understanding is still vulnerable to the nonidentity problem: since the agent would not exist had this specific action not occurred, there is no comparative action that the wrongdoer ought to have taken that would have made the agent any better-off. There is no other action that could have been done that would also result in the existence of the victim. So only the second understanding—when "worse-off than one ought to have been" is understood in terms of independent moral reasons—stands a chance to overcome the nonidentity objection. I consider this view below. For now, the point is that any comparative account of injury falls victim to the nonidentity objection in the same way as the counterfactual account, regardless of the baseline that is being compared.

So to avoid the same problem as any comparative account, we need an alternative account of injury that is *non*comparative. Let us consider two prominent examples of such accounts: the *bad state account* and the *threshold account*.[33] To

32 See Boonin, *The Non-Identity Problem and the Ethics of Future People*, 60–64.

33 For the bad state account, see, e.g., Harman, "Can We Harm and Benefit in Creating?"; and Johansonn and Risberg, "A Simple Analysis of Harm." For the threshold account, see, e.g., Meyer, "Past and Future" and "Intergenerational Justice." For an overview of theories of

challenge P2, we need to show that one of these accounts can do better than the counterfactual account in a way that generates claims to redress on the WIC model in cases of historical injustice.

The bad state account says that an action injures someone when it puts them in a bad state. For instance, on Elizabeth Harman's version, an action puts an agent in a bad state when the action "causes pain, early death, bodily damage, or deformity to [the agent]."[34] Crucially, an agent can be in a bad state even if that agent is also benefitted by the action.[35] So the bad state account seems to overcome the nonidentity problem. Agents can be put into a bad state by actions even if they would not exist without these actions.

However, the bad state account cannot be used by the WIC model in a way that plausibly generates claims to redress. Here is why. Any version of the WIC model says that experiencing an injury is a necessary condition to generate a claim to redress. However, if injury is defined in terms of the sorts of bad states that Harman specifies (pain, early death, bodily damage, or deformity), then this model undergenerates claims to redress if used in a theory of redress. Consider some basic wrongs like property theft. If I steal your bicycle and use it while you are away on vacation, I wrong you and seem to owe you redress. Yet I do not necessarily put you in pain, cause an early death, or so on. Moreover, imagine cases in which an agent obviously wrongs someone (e.g., assault) but due to the victim's neurological condition, do not happen to cause pain (or deform, etc.). Assault obviously requires redress even if it does not put the victim in a bad state. So these basic bad states are obviously bad yet seem to undergenerate claims to redress.

One might try to amend the bad state account to overcome this shortcoming by expanding the list of what counts as bad states (and therefore injuries). Perhaps not only these extreme states but *any state that goes against a victim's*

---

harm, see also Purves, "Harming as Making Worse Off"; Johansson and Risberg, "A Simple Analysis of Harm"; Carlson et al., "Plural Harm." A kind of account I do not consider here that is similar to the bad state account is Pitcovski's *intrinsic explanation account* ("Explaining Harm"). This account says (roughly) that an event injures an agent when the totality of states that can be explained by this event are intrinsically bad for the agent. However plausible Pitcovski's account may be, it does not help the WIC model in the case of the nonidentity problem. The nonidentity problem assumes that agents live lives worth living. By assumption, then, the event that causes their birth explains a totality of states that are intrinsically good for the agent. So the agent cannot be said to be injured by the event. This problem applies to any similar "causation" account of injury when applied to the WIC model in order to overcome the nonidentity problem. See Pitcovski, "Explaining Harm," for the similarity between his account and causation accounts.

34 Harman, "Can We Harm and Benefit in Creating?" 93.

35 Harman, "Can We Harm and Benefit in Creating?" 91–95.

interests might count as bad states (and therefore injuries). However, if we expand what counts as a bad state in this way, then if used by the WIC model, this account *overgenerates* claims to redress. Imagine:

> *Low-Cost Rescue*: Suppose you are drowning in your backyard pond. I am walking by your house and hear you. I trespass on your lawn and rush to save you. I save your life but trample your prizewinning manicured lawn in the process.

In Low-Cost Rescue, I save your life (benefit you) and yet go against your interest in your prizewinning manicured lawn. If the WIC model incorporated the expanded bad state account, I would owe you redress for my action because it goes against your interest. This seems counterintuitive.

The crux of the issue with the bad state account is this. Harman restricts what counts as a bad state to extreme states (pain, deformity, death, and so on) to avoid tricky cases like nonidentity cases. If we keep the account restricted to such extreme cases, the account undergenerates in the WIC model as a general theory of redress. If we expand the idea of bad state beyond its intended set of cases, however, then we have an unintuitive account of injury that *over*generates claims to redress. So the bad state account cannot overcome the nonidentity objection using the WIC model since it cannot satisfyingly generate claims to redress.[36]

A second alternative account of injury is the *threshold account*, developed by Lukas Meyer.[37] The threshold account says that an action injures an agent iff that agent is made worse-off than they ought to have been, where "ought to have been" is understood in terms of a *noncomparative* well-being threshold. The threshold account needs to specify this threshold nonarbitrarily. Meyer develops the threshold account to overcome nonidentity worries in the context of obligations to future generations. So he specifies the threshold using a sufficientarian standard: the well-being threshold is set at whatever is required to enjoy a minimally good life.[38] When an action makes an agent drop below this threshold required for a minimally good life, the action injures that agent.

36 It might be pointed out that the unintuitive result in this example (saying that you are injured by my rescuing you) could be avoided if we consider the bad state as the conjunction of the lawn being trampled *and* being alive. This should not be considered a bad state when compared to the relevant counterfactual state in question—namely, the conjunctive state of your lawn not being tramped *and* being dead. If we take the relevant state to be the conjunctive and compare it to the relevant counterfactual, then the bad state account does not unintuitively overgenerate. However, to do this, we need to turn it into a comparative account. Thanks to an anonymous reviewer for encouraging me to point this out.

37 See Meyer, "Past and Future" and "Intergenerational Justice."

38 Meyer, "Intergenerational Justice."

As before, my goal is not to assess the threshold account as a general account of injury nor to consider whether it successfully overcomes the nonidentity problem generally. However, like the bad state account, the threshold account cannot provide a plausible alternative to the WIC model when it comes to generating claims to redress. Indeed, it creates the same result as the nonidentity objection for cases of historical injustice and says that redress for many historical injustices is not possible.

The threshold account's problem is as follows. Many wrongs and injuries that intuitively require redress do not count as injuries on the threshold account. Again, consider a simple bicycle theft. Intuitively, this is a wrong and results in an injury. This theft, however, does not drop the victim below the threshold of well-being sufficient to achieve a minimally good life. The threshold account would not say that the victim is injured by the wrong. To address this worry, one might amend the account to use a different threshold. But notice that the problem persists regardless of where the account specifies the threshold. No matter where the threshold is, we can find a case in which someone is made worse-off generally but not worse-off at the threshold at which something counts as an injury. This problem persists until we set the threshold low enough that any action that goes against the agent's interests counts as making that agent worse-off than the agent ought to have been. But then the threshold account collapses into the expanded version of the bad state account above. And as we just saw, this expanded idea of injury overgenerates claims to redress.

The threshold account faces a second problem. Even if the threshold account overcomes the nonidentity problem in some cases of intergenerational justice, it does not necessarily overcome the nonidentity objection to redress for historical injustices. If present agents do not fall below the relevant threshold, then they are not injured and thus do not have claims to redress. However, on this picture, many historical injustices do not require redress because they do not injure anyone in the present. Imagine, for instance, descendants of Jewish victims of pogroms who were forced to flee to North America. Many of these descendants now enjoy lives above a minimally good threshold. The threshold account says that these descendants are not injured by the historical injustice and thus have no claims to redress. Similar arguments apply to any other historical injustice (including, for instance, land thefts or slavery) if the present agents have well-being sufficiently above the minimal threshold. However, this seems to be a counterintuitive result and rules out many of the claims to redress in the same way the nonidentity objection does.[39]

39 Notice that this problem also arises for the bad state account. If present agents are not in pain, deformed, etc. because of the historical injustice, then present agents are not injured by it and cannot have a claim to redress. No doubt, the past victims of historical injustice

So even if there are alternative accounts of injury, we cannot appeal to them in order to reject P2. In cases of historical injustice, none of these accounts result in us saying that present agents are injured by the historical injustice. They either result in the same problem that individuals would not exist and are thus not injured. Or they result in us saying that agents are not injured because they are not in a sufficiently bad state or under a relevant threshold of well-being. This means that any account of injury we use faces the same result as the counterfactual account: present individuals whose existence depends on the historical injustice are not owed redress for the historical injustice. More importantly, these alternative theories are not compatible with the WIC model to create general theories of redress. So on any account of injury that is compatible with the WIC model, P2 holds, and no alternative account of injury can overcome the nonidentity objection.

4.1.3. Defense of P3

P3 says that for all present agents and most (if not all) historical injustices, historical injustices are necessary conditions for these agents having come to exist.[40] Whoever is alive in the present is a function of their parents having met and conceived them at a specific time. Historical injustices are so large-scale, affect so many people, and have dramatically changed the course of history such that they are necessary conditions for the particular people who are alive in the present having come to exist. It is impossible that the parents of the people alive in the present would have met and conceived their children at the same time if the injustice had not occurred.

To illustrate, let us again take the example of chattel slavery. The slave trade was such a monumental series of events that shaped the world profoundly such that present individuals would not exist had the injustices not occurred. Yet any individuals in the present we might take to be injured by the slave trade—say, Black Americans—exist only because the transatlantic slave trade happened. The web of causal effects is such that had slavery not occurred, no individual alive would have been conceived at the precise moment they were such that *this particular individual* exists. This point applies to anyone who might claim to be injured by any historical injustice. For any individual in the present, their ancestors would not have met and conceived their children at the time they did had the historical injustice not occurred.

---

were put into bad states by the injustice. But not all present agents who plausibly have claims to redress are in these sorts of bad states.

40 For a detailed argument for this point, see Sher, "Transgenerational Compensation."

4.1.4. Defense of P4

P4 follows by *modus ponens*. If an event is a necessary condition for an agent coming to exist (P2), then this agent cannot be said to be injured by this event, and for any historical injustice, that historical injustice is a necessary condition for any present agents having come to exist (P3), then, no present agent is injured by any historical injustice. P4 does not deny that *past* individuals are injured by historical injustice. And P4 does not rule out that subsequent wrongs related to the wrongs of the historical injustice committed after the actions of the historical injustice may injure present agents. This may allow for a "chain" of wrongs leading from the past injustice to the present, the most recent of which injures present individuals.[41] All P4 says is that the *past actions* of the historical injustice do not injure any present agents.

4.1.5. Conclusion and Challenge to the WIC Model

The conclusion follows from *modus ponens*. No present agent experiences injuries as a result of a historical injustice (P4). So if experiencing an injury is a necessary condition for a present agent to have a claim to redress (P1), no present agent has a claim to redress for a past action of historical injustice.

The conclusion to the nonidentity objection, then, challenges the WIC model. We want a model of redress for historical injustice that generates claims to redress. And if we want to overcome the supersession thesis with this model, we need it to generate redress based on there being agents who experience injuries so we can apply the strategy of distinguishing between restitution and reparation for these injuries. However, the nonidentity objection challenges the WIC model. No present agents experience injuries, so no present agents have claims to redress. Thus, the WIC model cannot overcome the supersession thesis by saying that claims to reparations for injuries avoid the supersession thesis. Present agents are not injured by the historical injustice and so cannot have claims to reparations.

I will now consider a second independent objection that challenges the WIC model.

4.2. The Causal Problem

On the WIC model, a necessary condition of any agent having a claim to redress is that that agent experiences an injury that is *the result of* a wrongful action ϕ. This requirement that the injury is the result of ϕ creates a *causal problem*. Here

41 See Boxill, "A Lockean Argument for Black Reparations"; Sher, "Transgenerational Compensation"; and Butt, "Nations, Overlapping Generations, and Historic Injustice."

is the general structure of the problem.[42] Between the past wrongs committed at the time of the historical injustice and the present, there have been nearly infinite actions, potential actions, or omissions. Each of these actions, potential actions, and omissions might have produced injury I that present agent A experiences. Put another way, because of the amount of time that has passed between the historical injustice and the present, any present injury I is *overdetermined*. Accordingly, we cannot say that the wrong of the historical injustice φ is the *cause* of the present injury I that generates claims to redress. Some other action between the past and the present could have also been the cause of I or would have caused I even if φ had not. And thus, many cases of historical injustice cannot meet the necessary condition of the WIC model.

This version of the causal problem relies on a *counterfactual theory* or "difference making" theory of causation. On the counterfactual theory, cause c causes effect e when, had c not occurred, e would not have occurred. So on the counterfactual theory, action ϕ causes I when, had B not committed ϕ, A would not have experienced I. The counterfactual theory is philosophically popular.[43] On the counterfactual theory, the causal problem is clear. There are so many intermediate actions, potential actions, or omissions between the historical injustice ϕ and the present. Each of these potential actions *could have* produced I. That is, there are so many other potential causes that could have lowered A's well-being in a similar way to ϕ; A in the present might have experienced I anyway had B not done ϕ. Take the example of chattel slavery in the United States. Suppose that a present Black individual A experiences a bad state that would intuitively count as an injury I (e.g., lower well-being than white individuals, lower well-being than A ought to experience, a state that goes against A's interests, or the like), and we want to say that I is the result of slavery such that A is owed redress for slavery. However, there are many potential actions that did and could have occurred between the time of chattel slavery and the present such that A would have experienced I anyway. For instance, the government of the United States might have instituted some other policy that would have injured A (e.g., Jim Crow laws, racist redlining policies, etc.). So according to

42 For versions of this problem, see Waldron, "Superseding Historic Injustice," 5–8; Sher, "Ancient Wrongs and Modern Rights," 13; Kershnar, "Are the Descendants of Slaves Owed Compensation for Slavery?"; and Boxill, "A Lockean Argument for Black Reparations," 66–67.

43 See, e.g., Lewis, "Causation"; Sher, "Ancient Wrongs and Modern Rights"; and Boxill, "A Lockean Argument for Black Reparations."

the counterfactual theory, we cannot say that ϕ is the cause of I since it does not make a difference to whether or not I came about.[44]

While popular, the counterfactual theory faces some well-known problems.[45] One might therefore object that we can avoid the causal problem if we adopt an alternative theory of causation. One prominent alternative to counterfactual theories are *regularity theories of causation*.[46] On regularity theories, c causes e when e is the sort of thing that regularly occurs with c. A sophisticated regularity theory is Richard Wright's theory of NESS causation.[47] On NESS causation, c causes e when c is a *necessary element of a sufficient set* that could cause e. A sufficient set is a set of events that, if they together occur, e also occurs. The difference between NESS causation and the counterfactual theory is that NESS causation says that *any element in any sufficient set* is a cause. But we do not need to say that without c, e would not have occurred for us to say that c is a cause of e. If c causes e in the way regularity theories specify, then c is a "causal contribution" to e.[48]

Regularity theories might seem to naturally solve the causal problem for cases of historical injustice. While ϕ does not make a difference to bringing about I, in most cases of historical injustice, ϕ *is* involved in a *sufficient* cause to bring about I. Accordingly, on a regularity theory of causation such as NESS causation, historical injustice ϕ causes an injury I that present agent A experiences. For instance, while chattel slavery was not the *only* possible cause of present Black individuals' lower well-being, it was a necessary element of a set *sufficient* to bring this lower well-being about.

However, regularity theories face a problem when they are incorporated into the WIC model and used to explain redress for historical injustice. As Frank Hindriks argues in the context of overdetermination and the problem of collective harm, regularity theories overgenerate responsibility.[49] Suppose the WIC

44 The causal problem is a metaphysical problem about whether historical injustice ϕ is the cause of injury I. It is thus not an epistemological question about whether we can *know* whether ϕ is the cause of I. Thanks to an anonymous associate editor of this journal for pointing this out.

45 See Paul and Hall, *Causation*.

46 See Mackie, *The Cement of the Universe*; and Wright, "Causation in Tort Law." Mackie's INUS theory is very similar to Wright's theory that I consider here, but I restrict my discussion here to Wright's NESS theory due to space constraints.

47 Wright, "Causation in Tort Law."

48 See Nefsky, "Collective Harm and the Inefficacy Problem"; and Hindriks, "The Problem of Collective Harm."

49 Hindriks, "The Problem of Collective Harm," 214–25. Thanks to an associate editor for the helpful suggestion to consider overdetermination and collective harm. There is an important difference between overdetermination in the collective harm context and cases

model adopts regularity causation as its account of causation. This means that *A* would be owed redress for ϕ iff ϕ is wrong and *A* experiences an injury *I* where ϕ is a causal contribution to *I*, understood as ϕ being a necessary element of a sufficient set, though not necessarily a difference maker, to produce ϕ. However, wrongs may causally contribute to injuries in ways that intuitively should not result in claims to redress for these injuries. Consider the following example:

> *Bicycle and Bahamas:* Suppose *A* steals *B*'s bicycle, which was *B*'s primary form of transportation. *B* must now take the bus to work and catches a cold. *B* spreads this cold to colleague *C*. *C* was scheduled to go on a vacation to the Bahamas, where *C* would have stayed in a resort owned by *D*. *C* is forced to cancel the vacation, and *D* loses income. *D* lashes out at their partner *E*, who experiences psychological harms.

In this example, *A*'s bicycle theft is wrong. And this wrong is a necessary element of a sufficient set to cause the injuries that *D* and *E* experience. However, it seems incorrect to say that *A* owes *D* or *E* redress for the injuries. Note further that we cannot even avoid this problem by insisting that someone other than *A* owes redress to *D* and *E*. Many injuries we experience have wrongs that are causal contributors. For instance, when I stub my toe on my desk, this is something for which (*ceteris paribus*) no one owes me redress. This is true even if it turns out that the person who sold me the desk stole the wood they used to make this desk. Even though a necessary element of a sufficient set needed to result in my injury is a wrong, it seems intuitively incorrect to say that I am owed redress for this injury. The more general point is that wrongful actions often causally contribute to injuries that are intuitively distant from the wrongs.[50] If the WIC model took causal contribution (regularity) as its version of cause, then we would implausibly overgenerate claims to redress.

---

of redress I consider here. In collective harm cases, part of what is at issue is whether the contribution to the collective harm makes the action wrongful. (See Nefky, "Collective Harm and the Inefficacy Problem.") The idea is that in such cases, the action would be wrong only if it contributes to the collective harm. So then the challenge becomes determining if the agent did contribute to the collective harm. However, this is different from how overdetermination plays a role in redress. In cases of redress, we already *assume* that the action is wrong. The question that arises is whether this action *causes injury I*. At issue is not whether that action is wrong but whether that action *caused the injury* that present individuals experienced. So whereas in cases of collective harm, there is pressure to say that the action did contribute to the collective harm to explain why the action is wrong, this same pressure does not occur in cases of redress because we have already assumed that the action is wrong.

50 Hindriks, "The Problem of Collective Harm," 214–25.

It might be objected here that we must amend the WIC model. In Bicycle and Bahamas, *D* and *E* are not wronged by *A*'s theft. The theft is a wrong done to *B*. So perhaps we must say that the WIC model generates a claim for reparations iff ϕ results in *I*, *A* experiences *I*, ϕ is a wrong, *and* ϕ *wrongs* A. Then we would not say that *D* or *E* is owed redress, nor am I owed redress when I stub my toe.

However, this move does not avoid the overgeneration problem.[51] We can still imagine cases in which *A* is wronged by *B*, *A* experiences an injury as a result of the wrong in the NESS sense, but intuitively, this injury should not be among redress for *B*'s wrong against *A*. Consider the following case.

> *Trespass and Terrorism*: Suppose *B* trespasses and parks in *A*'s driveway, blocking *A*'s car from getting out. *A* must take the bus to get to work. This bus is targeted by a terrorist attack in which *A* sustains serious injuries that result in long-term physical disabilities and psychological trauma.

In this case, *A* is wronged by *B*, and this wrong is a causal contribution to *A* experiencing a significant injury *I*. However, it seems implausible that *B* should be responsible for redressing *A*'s injuries. *B* owes some redress for the trespass. But this redress seems not to be needed for the injuries sustained in the terrorist attack. An account that requires the latter overgenerates redress. So mere causal contribution as specified by regularity theories like NESS does not seem to help the WIC model.

My aim here is not to adjudicate between different theories of causation. Rather, I consider these versions of causation to point to a common problem with the WIC model when it comes to redress for historical injustice. Different theories of causation specify different conditions under which we can say that injury *I* is the result of historical injustice ϕ. However, regardless of the theory of causation that the WIC model incorporates, there is a common problem. There is a clear wrong (ϕ). But the WIC model says that the claim to reparations is not generated by ϕ. Rather, it is generated by being able to draw a causal line from that wrong to some injury *I*. But this means that once it is difficult to draw a clear causal line from ϕ to *I*, the possibility of redress is called into question even though ϕ occurred and was wrong. This is counterintuitive. Put another way, regardless of the theory of causation that the WIC model incorporates, it seems that the possibility of redress can always be called into question simply by raising worries about the causal connection between ϕ and *I*. However, this seems counterintuitive when we are certain that ϕ occurred, and present agent *A* is a victim of ϕ. Something is going wrong here. The possibility of reparations should not be able to be called into question simply by questioning the relation

51 For a similar argument, see Perry, "The Moral Foundations of Tort Law," 462–67.

of the wrong to certain injuries. However, since the WIC model makes the fact that ϕ results in injuries a necessary condition of redress, the WIC model allows for redress to be challenged in this way. This suggests we should turn to a different theory of redress that cannot be challenged in the same way.

4.3. Response: The Structural Turn

One might try to resist the causal problem by appealing to *structural redress*.[52] Perhaps we can accept that any injury a present individual experiences is not the causal result of any particular discrete historical injustice. However, we can still insist that present structures are the result of the aggregate of historical injustices. For instance, racial inequality and injustice in the United States would not have occurred had all the historical injustices involved in slavery and Jim Crow not occurred. Proponents of *structural reparations* argue that reparations are about addressing the unjust structures that have resulted from historical injustice.[53] We have reasons of structural or distributive justice to address unjust structures, and so there are claims to redress historical injustices when historical injustices have created unjust structures. This avoids the causal problem. We can insist that present unjust structures would not have resulted had the aggregate of historical injustices not occurred. And since these unjust structures require repair, there are claims to redress even if the causal problem challenges the possibility of redress for discrete injustices.

The structural turn is an important development in the historical injustice literature. I have defended its importance in other work.[54] However, appealing to it does not solve the causal problem that the WIC model faces when it comes to overcoming the supersession thesis. While structural reparations might generate redress for historical injustice, they do not generate the kind of redress we are trying to vindicate. Recall that the supersession thesis challenges the possibility of redress for particular injustices. In Waldron's original example, for instance, the supersession thesis challenges the possibility of redress for *B*'s land expropriation from *A*. The supersession thesis therefore challenges the possibility of *interactional redress*. Interactional redress concerns what one agent is owed because that agent has been a victim of an injustice. Interactional

52 Thanks to an anonymous reviewer for encouraging me to consider the structural turn here.

53 See Lu, *Justice and Reconciliation in World Politics*; Nuti, *Injustice and the Reproduction of History*; Song, "Denial of Japan's Military Sexual Slavery and Responsibility for Epistemic Amends"; Butt, "What Structural Injustice Theory Leaves Out"; and Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice."

54 Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Reparations and Structural Injustice."

redress is different than structural reparations.[55] Structural reparations concern what is required by justice to ameliorate or repair unjust structures and unjust distributions for reasons of structural justice. This means that structural reparations for historical injustice are required for reasons of justice that appeal to the *structure* of past wrongs, not the fact of the past wrongs. Where interactional redress is owed *for the past wrong*, structural reparations are owed *for the fact that the structure is unjust*.[56] For this reason, interactional redress is often referred to as backward looking, while structural reparations are often referred to as forward looking.[57] In many cases, the *cause* of the unjust structure is the historical injustice. But the *normative reason* that generates redress is the unjust structure. The unjust structure requires repair for forward-looking reasons regardless of what created it. In contrast, in interactional redress, the reasons that generate redress are the past wrongs themselves, for backward-looking reasons. Thus, structural reparations and interactional redress are different.

This difference is significant when it comes to using the structural turn to overcome the causal problem. Suppose we say that redress is owed for reasons of structural justice because the historical injustice resulted in the unjust structure. This vindicates the possibility of structural reparations. However, this still does not vindicate the possibility of interactional redress. This is significant for two reasons. First, the supersession thesis challenges the possibility of interactional redress. And so the target we are trying to vindicate here is the possibility of redress for the interaction itself. Structural reparations do not vindicate this; they vindicate the possibility of *some* redress, but not of the particular redress that the supersession thesis challenges. Any present agent would have a claim that the unjust structure they are disadvantaged by ought to be repaired but *not* a claim to the particular wrong that may have wronged or injured them being repaired. So structural reparations do not vindicate the kind of redress that the supersession thesis challenges. Moreover, not only do structural reparations struggle to vindicate interactional redress, but they might actually *challenge* the possibility of redress in a very similar way to the supersession thesis.

The supersession thesis challenges the possibility of redress for a particular past wrong when reasons of present justice supersede redressing the past wrong. The structural approach to historical injustice might do something similar. Structural reparations require repairing an unjust structure for reasons of

55 Lu, *Justice and Reconciliation in World Politics*; Nuti, *Injustice and the Reproduction of History*; and Butt, "What Structural Injustice Theory Leaves Out."

56 Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Reparations and Structural Injustice."

57 Song, "Denial of Japan's Military Sexual Slavery and Responsibility for Epistemic Amends"; and Butt, "What Structural Injustice Theory Leaves Out."

structural justice. If the *only* reparations that were required were structural reparations, then structural reparations might actually say that discrete wrongful interactions ought *not* be repaired for reasons of structural justice. For instance, Alasia Nuti argues that addressing discrete interactional wrongs is redundant and may even distract from the more important structural concerns.[58] Accordingly, redress for discrete wrongs that are truly in the past, she argues, ought to sometimes be ignored for reasons of present structural justice.[59] Catherine Lu makes similar arguments in her defense of the structural model over the interactional model.[60] This starts to look like the challenge that the supersession thesis poses for redress. There are reasons of structural (present) justice that might supersede the reasons of backward-looking interactional justice to address a past wrong. So appealing to reasons of structural justice might not vindicate the kind of interactional redress from the supersession thesis in the way we aim to do here and might even pose a similar problem as the supersession thesis.

Second, structural reparations and interactional redress might differ in terms of the content of reparations. Accordingly, even if appealing to the structural model generates *some* redress, something is lost if structural reparations are all we have. One important difference concerns who owes the reparations. Many have argued that interactional reparations are owed by the wrongdoer of the injustice, while structural reparations are owed by every member of society.[61] In many cases, there is a powerful intuition that we want redress for past wrongs to come from the particular wrongdoers of the injustice. This seems to require vindicating interactional redress. Similarly, because structural reparations are owed to improve the structure, they might not be particularized to the historical injustices that causally brought about the structure. Suppose structure *S* is unjust because individuals *A, B, and C* experience an insufficient amount of good *x*. Suppose structure *S* was brought about by a historical injustice. Structural reparations would require repairing *S* to bring *A, B,* and *C* up to a sufficient amount of good *x*. But good *x* might have nothing to do with the historical injustice; the historical injustice might have been wrong because it interfered with a different good *y*. In other words, nothing in the structural reparations would be *particularized* to the historical injustice. However, in many cases, it seems that a desideratum of reparations is that reparative measures

58 Nuti, *Injustice and the Reproduction of History*, 15–16, 157.

59 Nuti, *Injustice and the Reproduction of History*, 51.

60 Lu, *Justice and Reconciliation in World Politics*, 114–40.

61 See Weinrib, *The Idea of Private Law*; Ripstein, *Private Wrongs*; and Lu, *Justice and Reconciliation in World Politics*.

reflect the particular reasons that an action was wrong.[62] Finally, there might be kinds of redress that are primarily backward rather than forward looking. For instance, apologies and acknowledgements of wronging are often among the most demanded and symbolically important elements of reparations for a wrong. These practices are importantly backward looking.[63] Even if they also have a forward-looking purpose, it seems that *part* of their goal is backward looking and concerns what the wrongdoer did rather than what would be good for the society going forward.[64] If we had only structural reparations for historical injustice, then we might lose the important things that interactional redress provides that structural reparations might struggle to.

For these reasons, others and I argue that reparations for historical injustice must involve *both* structural and interactional elements.[65] To be clear, my argument here is not that structural accounts of reparations are not helpful. Indeed, in other work I argue extensively that they are indispensable.[66] Nor do I deny that historical injustices have resulted in unjust structures. I think they have. My point here is that even if structural reparations are also required, they do not achieve all the same things that interactional reparations do. An optimal account of redress for historical injustice explains how both structural and interactional redress is required. The supersession thesis and the causal problem for the WIC model challenge the possibility of interactional redress. So if we want to vindicate an optimal account that includes both interactional and structural redress, we still need to overcome the causal problem.[67]

62 Ripstein, *Private Wrongs*, 125; and Weinrib, *The Idea of Private Law*, 60–90.

63 Song, "Denial of Japan's Military Sexual Slavery and Responsibility for Epistemic Amends."

64 Helmreich, "The Apologetic Stance"; Song, "Denial of Japan's Military Sexual Slavery and Responsibility for Epistemic Amends"; McKeown, "Backward-Looking Reparations and Structural Injustice"; and Lambrecht "Reparations and Structural Injustice."

65 Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Reparations and Structural Injustice"; McKeown, "Backward-Looking Reparations and Structural Injustice"; and Butt, "Nations, Overlapping Generations, and Historic Injustice."

66 Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Reparations and Structural Injustice."

67 There is a second way someone might appeal to structures to overcome the causal problem. Perhaps historical injustices (in the aggregate) caused present unjust structures, and these unjust structures disadvantage some present agents. We might say that these present agents the unjust structure disadvantages are injured by the historical injustice. For reasons of backward-looking justice, these present agents might be owed redress for the historical injustices since they experience an injury as a result of the injustices in the way the WIC model requires. This might be an *interactional* version of a structural model: redress is owed because present agents who are disadvantaged by the unjust structure are injured by the historical injustice and, on these grounds, are owed redress for the historical

5. The Wrongful Action Claim Model to the Rescue

The WAC model, the other model of redress, can overcome the supersession thesis without encountering the problems that the WIC model faces. I begin this section by presenting the WAC model. I then show how it avoids the nonidentity and casual problems. Finally, I demonstrate how the WAC model can incorporate Harrison's insights about the distinction between restitution and reparation to overcome the supersession thesis.

My goal here is to demonstrate that the WAC model can overcome the supersession thesis without facing the nonidentity and causal problems that the WIC model faces. The WAC model is widely developed outside the context of historical injustice.[68] Others develop the WAC model in the context of historical injustice.[69] I also develop the model in detail elsewhere.[70] Therefore, my purpose here is not to develop this model in detail. Rather, my purpose is to demonstrate that we ought to prefer it over the WIC model because it can overcome the supersession thesis and the problems that the WIC model faces. Accordingly, I present the WAC model only as needed for this purpose.

---

injustice. However, this argument does not vindicate the WIC model. This is for two reasons. First, the same causal problem applies here. While we may unquestioningly say that the historical injustices together caused the unjust structure, it is not clear that they caused any present agent's injuries. Even if we allow with certainty that the historical injustices in the aggregate were the cause of the unjust structure, this does not show that the historical injustices were the cause of the contemporary agent's injury. The structure might have been just one way that the injury would have come about. So the historical injustices (by way of the structure) might not be the *difference-making* cause of any present agent's injury. Once again, the general causal problem shows up: once we can question whether a present injury really is the result of historical injustices, we can question the possibility of (interactional) redress. As I argue above, this is counterintuitive. If the present agent is wronged by the injustice, the causal line from the injustice to the injury should not make us question the possibility of redress. Second, this argument encounters the nonidentity problem once again. To be interactional rather than structural, this argument needs to say that present agents are injured by the historical injustice and on these grounds are owed reparations. However, as we have seen, the nonidentity problem challenges this claim. Present individuals cannot be injured by the historical injustice because without the historical injustice, they would not exist. Even if this injury comes by way of the structure, if the reasons that generate reparations are the backward-looking injuries themselves, then the nonidentity problem challenges the argument.

68 E.g., Ripstein, *Private Wrongs*; Weinrib, *The Idea of Private Law*; and Gardner "What Is Tort Law For?"

69 E.g., Kumar, "Why Reparations?"; Shiffrin, "Reparations for US Slavery and Justice over Time"; Thompson, "Historical Injustice and Reparation"; and Sanderson, "Redressing the Right Wrong."

70 Lambrecht "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

5.1. Presenting the WAC *Model*

Recall the WAC model:

> Agent *A* has a claim to redress for wrongful action ϕ iff *A* is wronged by ϕ.

The most important difference between the WAC and WIC models is that, unlike the WIC model, the WAC model does not make experiencing an injury a necessary condition of generating a claim to redress. Instead, the WAC model says that the fact that a victim was wronged is sufficient to generate a claim. There is variation between versions of the WAC model in what counts as a wrongful action. Some versions focus on rights violations.[71] Others take a contractualist approach and say that *B* wrongs *A* when *B* does an action that violates *A*'s normative expectations.[72] The differences between what counts as a wrong do not matter for my purposes here. On any plausible theory of a wrong, the actions involved in historical injustices are certainly wrongs. Moreover, as we have seen, an action need not result in an injury to be a wrong. Of course, the WAC model does not deny that wrongs often also involve injuries. But even when they do not, the *normatively significant feature* that makes an action wrong is what generates the claim to redress, *independent* of any injuries that result from the wrong. This means that what generates a claim to reparations is the fact that the action violates a right, expresses something about the victim, violates contractualist normative expectations, or violates a duty the wrongdoer has to the victim for some other reason.

The WAC model determines what redress consists of by looking at these normatively significant features of the wrong.[73] The normatively significant features of the wrong correspond to the moral reasons the wrongdoer ought not to have done the action to the victim.[74] Redress can include *material compensation* and *moral repair.* Let us look first at compensation. Consider a simple bicycle theft. The reason the theft is wrongful is because I interfered with an object of your property that should have been yours to control. Redress for this theft requires reestablishing the situation that ought to have occurred in which you can use the property as you see fit. This most straightforwardly includes the return of the bicycle. But it might also include compensating you for losses that you sustained while without the bicycle. If I do not compensate those losses

71 E.g., Ripstein, *Private Wrongs*; Weinrib, *The Idea of Private Law*; and Gardner, "What Is Tort Law For?"

72 Kumar, "Why Reparations?"

73 Kumar, "Why Reparations?"; and Thompson, "Historical Injustice and Reparation," 118.

74 See Hurley and Weinberg, "Whose Problem Is Non-Identity?"

and you continue to bear them, I continue to interfere with what you can do with the property.

Those who develop the WAC model offer different ways we can determine material compensation owed for a wrong. One approach, for instance, is that material compensation is required to restore the means that were interfered with by the wrong.[75] Another approach is to say that any injury a victim experiences continues to interfere with them in the way the wrong originally did and so should be viewed as part of the wrong that reparations must repair.[76] Another approach argues that compensation ought to be determined by way of substantive moral argument based on the kind of injury that accompanies the wrong.[77] The crucial point is that the WAC model does not simply say that the precise good that was interfered with must be returned. Reparations likely also require material compensation. However, unlike the WIC model, the *reason* material compensation is owed is not because the victim experiences an injury that results from the wrong. Rather, for the WAC model, the reason compensation is owed is because this compensation is necessary to establish the situation that ought to have occurred. Fully establishing a situation in which a victim is not wrongly interfered with by a wrongdoer requires that the victim is also compensated for the ways the wrong interfered with her.

Adequate redress also often includes moral repair (such as apologies or acknowledgments of wrongdoing). Like the other elements of redress, moral repair responds to the normatively significant features that made the action wrongful by demonstrating that the wrongdoer ought not have treated the victim in the particular way the wrongdoer did.[78] The content of redress on the WAC model is determined by looking at whatever about the action made it wrongful, and it will thus vary based on the particular wrong in question. One popular way the WAC model might do this is by appealing to the *continuity thesis*.[79] The continuity thesis states that the reasons that generate redress are the same reasons that made the action wrong in the first place. For instance, *A* might violate *B*'s right by stealing *B*'s bicycle. The reasons that the action was wrong is because *A* violated *B*'s right. This means that reparations must now respond to the same reasons that made the action wrongful. *A* must now restore *B*'s right by restoring *B*'s control over the bicycle and compensating injuries that

75 Ripstein, *Private Wrongs*, 233–44.

76 Gardner, "What Is Tort Law For?"

77 Kumar, "Why Reparations?" 203.

78 Helmreich, "The Apologetic Stance."

79 Gardner, "What Is Tort Law For?"; Weinrib, *The Idea of Private Law*; and Ripstein, *Private Wrongs*.

resulted from the theft. The continuity thesis gives the WAC model a clear way to determine what reparations require by looking at the reasons the action was wrongful in the first place.

In the context of historical injustice, the WAC model must demonstrate how present agents can be wronged by historical injustices. There are various versions of this.[80] To illustrate the WAC model, I sketch two versions of an argument about how historical injustices wrong present individuals. My purpose here is not to endorse the specifics of either argument. Rather, my goal is to sketch different versions that can be used to demonstrate how the WAC model overcomes the problems that the WIC model faces.

First, consider an argument offered by Rahul Kumar about redress for the wrongs of slavery. Kumar argues that historical injustices not only wrong individuals alive at the time of the injustice but also wrong individuals not yet born. Past wrongs of the historical injustice wrong future individuals. When a past wrong is done against a "type" of person and when the present individual is a token of this type, the past wrong wrongs a present individual.[81] Kumar argues that one of the wrongs of chattel slavery was committed against Black Americans as a "type." In particular, the past actions of chattel slavery were (partly) wrong in that they treated Black individuals as having "inferior status" in "public reason."[82] At the time of the historical injustice, this inferiority was legally codified by chattel slavery laws. But because these wrongs treated every member of the type as inferior, these wrongs also wrong present Black individuals. Thus, present individuals are wronged by this past wrong. Redress for this wrong requires making Black individuals equal to others. This might be achieved by legislation that establishes that Black individuals have equal status. But as Kumar argues, redress also likely must include material compensation.[83] The WAC model requires material compensation when it is needed to restore the situation that ought to have occurred that the wrong prevented. Establishing Black individuals' equal moral status very likely requires material compensation. Years of inferior treatment have produced large material inequalities that reinforce unequal moral status. And so to establish equal moral status, reparations must include some material compensation. Unlike the WIC model, the reason that generates this material compensation is not that Black individuals experience injuries that results from the wrong. Rather, the reason is that they

80 See Kumar, "Why Reparations?"; Thompson, "Historical Injustice and Reparation"; Shiffrin, "Reparations for US Slavery and Justice over Time"; and Lambrecht, "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

81 Kumar, "Why Reparations?" 208.

82 Kumar, "Why Reparations?" 205, 209.

83 Kumar, "Why Reparations?" 210–11.

are wronged, and restoring the situation that ought to have occurred that this wrong disrupted requires material compensation.

A similar argument for how past injustices wrong present agents is offered by Douglas Sanderson. Sanderson argues that the injustices involved in colonial wrongs in North America wronged Indigenous peoples by violating their rights to develop their social and political institutions.[84] These actions continue into the present. For instance, actions that made Indigenous governance subordinate to federal governments of colonial states (as in the Canadian context in which Sanderson writes) started in the past and continue to make Indigenous governance subordinate. Plausibly, present Indigenous individuals continue to have a right to develop their social and political institutions. Accordingly, these past actions also wrong present Indigenous individuals by violating their right to develop their political institutions.

Again, my discussion here is not meant to endorse either of these arguments. The point is to illustrate how a past action may wrong present individuals. When an argument does this, it can say that present individuals are wronged without necessarily also needing to say that present individuals experience an injury that is the result of the wrong.

I now turn to demonstrating how the WAC model overcomes the problems that the WIC model faces.

5.2. *The Nonidentity and Causal Problems*

5.2.1. *The Nonidentity Problem*

Unlike the WIC model, the WAC model does not make experiencing an injury that results from a wrong a necessary condition of a claim to redress. Instead, when an agent is wronged, the agent has a claim to redress regardless of whether or not that agent is also injured as a result of the wrong. This allows the WAC model to avoid the nonidentity objection. We can accept that a historical injustice does not injure any present individuals. But present individuals can still have claims to redress. If present individuals are wronged by the historical injustice, they have claims to redress for it.[85] Recall Kumar's argument to see this. The historical injustices involved in slavery were wrong because they treated Black Americans as inferior and disrespected their fundamental equality as fellow agents. This action has this significance even if it just so happens to not result in any harms to Black Americans. Thus, we do not need to say that

84 Sanderson, "Redressing the Right Wrong," 126–35.

85 For more details on this argument, see Shiffrin, "Reparations for US Slavery and Justice over Time"; Kumar, "Why Reparations?"; and Lambrecht, "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

Black Americans experience injuries that they would not have had the wrong not occurred. Rather, we can simply say that the wrong treats them as inferior regardless of any effects they experience. And so present individuals have a claim to redress regardless of whether they experience any injuries as a result of the wrong.

The WAC model overcomes the nonidentity problem by insisting that present individuals are wronged by historical injustice. This is a wrong-based solution to the nonidentity problem. However, there is a main objection to wrong-based solutions to the nonidentity problem.[86] The objection runs as follows. A moral agent is not wronged if that moral agent waives a right that an action would normally violate. In some cases, it seems that an agent's right is not violated if that agent is *unable to* waive the right, and we can reasonably conclude that the agent *would have* waived the right. Consider an unconscious patient who cannot consent to a lifesaving surgery. The patient cannot waive his right to bodily integrity. But given that he has a prudential reason to waive his right, he *would* waive this right, and, it is argued, the surgeon does not violate his right by operating.[87] More generally, whenever an agent has a moral or prudential reason in favor of waiving a right, we can reasonably conclude that the agent *would* waive the right. And when an agent *would* waive the right, and the agent cannot actually do so, it is argued, we can conclude that the agent *does* waive the right. In nonidentity cases, individuals would be choosing between a situation in which they would have a right violated and a situation in which they would not come into existence at all.[88] Individuals have a strong prudential reason to waive their right. Since we can conclude that they *would* waive their right, we can conclude that they *do* waive their right. Therefore, individuals in nonidentity cases are not wronged by the action.

This hypothetical rights waiver argument attempts to challenge my argument that the WAC model avoids the nonidentity problem. I consider this argument in greater detail elsewhere.[89] However, for the purposes of this article, it suffices to reply in two ways.

First, we can question the rights waiver argument itself.[90] The point of rights is that it is up to the rights holder whether or not to waive the right. The rights waiver argument seems to ignore this. In particular, the move from "we could

86 See Parfit, *Reasons and Persons*, 375; and Boonin, *The Non-Identity Problem and the Ethics of Future People*, 120–24.

87 Boonin, *The Non-Identity Problem and the Ethics of Future People*, 123.

88 Boonin, *The Non-Identity Problem and the Ethics of Future People*, 122.

89 Lambrecht, "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

90 See Hurley and Weinberg, "Whose Problem Is Non-Identity?"

reasonably conclude that the agent would waive the right" to "so the agent *does* waive the right" seems to ignore that the whole point of rights is that they protect things that others are not permitted to decide for you.[91]

Second, even if the rights wavier argument does challenge wrong-based solutions to the nonidentity problem generally, it does not challenge wrong-based solutions to historical injustices *in particular*. Historical injustices are one particular kind of nonidentity case. Central to the rights waiver argument is that we can reasonably conclude that an agent would waive the right if the agent had strong moral and prudential reasons to do so. In cases of historical injustices, present individuals might have prudential reasons to waive their rights: had the action not been done, the present individual would not exist. However, present individuals have strong *moral reasons* against waiving their rights. The past actions of the historical injustice were terrible wrongs done not only to present individuals but also to the past individuals who were victims of the injustice at the time of the injustice. Present individuals have strong moral reasons not to consent to these actions being done since these actions wrong past individuals. So we cannot reasonably conclude that present individuals would waive their rights and thus cannot say that they *do* waive their rights.[92]

So in cases of historical injustice, the rights waiver argument does not apply. Accordingly, the main objection that threatens wrong-based solutions to the nonidentity problem does not threaten the WAC model in the context of historical injustices. And so the WAC model can overcome the nonidentity objection.

5.2.2. *The Causal Problem*

The causal problem challenges the WIC model because it says that the injuries that present individuals appear to experience might not be the result of the wrong. The WAC model does not have this problem. On the WAC model, a victim experiencing injuries that result from the wrong is not a necessary condition to generate redress for the historical injustice. So we can admit that any injuries that present individuals experience might be the result of something else. But if present individuals are wronged by the historical injustice, then present individuals have claims to redress. Again, consider Kumar's argument.

91 Note that even if you adopt an interest theory of rights, the hypothetical rights waiver argument still misunderstands rights. Even if rights function to further the interests of rights holders, this does not mean that if a particular instance of a right fails to further the interest of a rights holder, it ceases to be a right or that the rights holder automatically waives it. Rather, on the interest theory, rights (or particular kinds of rights) must *generally* promote rights holders' interests.

92 Boonin's version of the rights waiver argument (*The Non-Identity Problem and the Ethics of Future People*, 266–67) even admits the point I make here.

We do not need to say that any harms that Black Americans experience are the result of historical injustices. Instead, we can simply say that Black Americans are wronged by historical injustices since these injustices treated them as inferior. Accordingly, the possibility that any injury in the present is not the result of a historical injustice does not challenge the possibility of redress. This vindicates our intuition: we *know* that the historical injustice wronged a present agent. And so the mere fact that an injury's cause can be questioned should not challenge the possibility of redress.

5.3. *Supersession*

The WAC model avoids the problems that the WIC model faces. We can now incorporate Harrison's strategy to overcome the supersession thesis into the WAC model. This yields a model of redress that can overcome the supersession thesis while not being vulnerable to the nonidentity and causal problems. The WAC model can incorporate the same move Harrison makes by distinguishing between restitution and reparation as two forms of redress. While the supersession thesis might prevent restitution, it does not prevent reparation. The WAC model can incorporate this distinction as follows.

Step one is to determine whether there is a claim to redress at all. This is the step where the WIC model falters because of the nonidentity and causal problems. The WAC model succeeds where the WIC model fails. If present individuals are wronged by the historical injustice, then present individuals have a claim to redress.

Step two is to determine what *kind* of redress is possible. Either restitution or reparation is possible. In some cases, the supersession thesis might make it such that restitution is not possible. However, even in these cases, reparation may still be possible. Whether restitution or reparation, the WAC model determines the content of redress by the normatively significant features of the wrong. For instance, in a property theft, the normatively significant feature is the interference with the victim's right to use the property as the victim chooses. So redress must in some way address the interference with this right. Restitution might do this by returning the precise object that was stolen. However, even if this is not possible, reparation can still address the normatively significant feature of the interference with the right. For instance, reparation might include compensation for the right being violated.[93] More importantly, it might include restoring the victim's ability to exclude the wrongdoer from interfering with the victim's rights *generally*, even if they cannot return the precise object that was stolen because of the supersession thesis. There are plenty

93 Ripstein, *Private Wrongs*, 125–45.

of arguments for how reparations can respond to the normatively significant features of a wrong in this way without returning the precise content of a wrong. For instance, Sanderson provides a compelling argument for how restoring the right like this might be possible in cases of Indigenous land expropriations. Sanderson argues that land expropriations violated Indigenous peoples' right to self-determination through political institutions. So even if returning the land (restitution) is not possible because of the supersession thesis, restoring the ability to develop political institutions may still be possible.[94] A. J. Simmons provides a similar argument by distinguishing between returning the precise object of a right and the right itself.[95] Finally, Harrison points to a number of possible reparative measures that are not restitution but would address the normatively significant features of the wrong.[96] My goal here is not to develop the way that the WAC model determines the content of redress or reparation. I, along with others, do this elsewhere.[97] Instead, my goal here is to show that the WAC model can apply the same strategy of distinguishing between restitution and reparation. This can overcome the supersession thesis. Even if the supersession thesis challenges restitution, it may not challenge reparation. Thus, on the WAC model, the supersession thesis does not challenge the possibility of claims to redress for historical injustices.

6. Conclusion

In this article, I have set out to develop an account of redress for historical injustice that is not vulnerable to the supersession thesis. By distinguishing between two forms of redress—restitution and reparations—an account of redress can overcome the supersession thesis. However, I have argued that a model that says claims to redress are generated by injuries resulting from the historical injustice (the WIC model) is vulnerable to the nonidentity and causal problems. In contrast, a model that says claims to redress are generated by the wrong of the historical injustice (the WAC model) does not face these problems.

94 Sanderson, "Redressing the Right Wrong."

95 Simmons, "Historical Rights and Fair Shares."

96 Harrison, "Supersession, Reparations, and Restitution."

97 Kumar, "Why Reparations?"; Sanderson, "Redressing the Right Wrong"; Simmons, "Historical Rights and Fair Shares"; Gardner, "What Is Tort Law For?"; Weinrib, *The Idea of Private Law*; Ripstein, *Private Wrongs*; and Lambrecht, "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice" and "Reparative Justice, Historical Injustice, and the Nonidentity Problem."

This means that if we want an account of redress for historical injustice that can overcome the supersession thesis, we must adopt the WAC model.[98]

Ludwig-Maximilians-Universität München
felix.lambrecht@lmu.de

REFERENCES

Anderson, Elizabeth S., and Richard H. Pildes. "Expressive Theories of Law: A General Restatement." *University of Pennsylvania Law Review* 148, no. 5 (2000): 1503–75.

Boonin, David. *The Non-Identity Problem and the Ethics of Future People*. Oxford University Press, 2014.

Boxill, Bernard. "A Lockean Argument for Black Reparations." *Journal of Ethics* 7, no. 1 (2003): 63–91.

Boxill, Bernard, and J. Angelo Corlett. "Black Reparations." *Stanford Encyclopedia of Philosophy* (Spring 2022). https://plato.stanford.edu/archives/spr2022/entries/black-reparations/.

Butt, Daniel. "Nations, Overlapping Generations, and Historic Injustice." *American Philosophical Quarterly* 43, no. 4 (2006): 357–67.

———. "What Structural Injustice Theory Leaves Out." *Ethical Theory and Moral Practice* 24, no. 5 (2021): 1161–75.

Carlson, Erik, Jens Johansson, and Olle Risberg. "Plural Harm: Plural Problems." *Philosophical Studies* 180, no. 2 (2023): 553–65.

Christie, Gordon. "The Supersession of Indigenous Understandings of Justice and Morals." *Critical Review of International Social and Political Philosophy* 25, no. 3 (2022): 427–42.

Feinberg, Joel. *Harmless Wrongdoing*. Oxford University Press, 1990.

Gardner, John. "What Is Tort Law For? Part 1. The Place of Corrective Justice." *Law and Philosophy* 30, no. 1 (2011): 1–50.

Goodin, Robert E. "Theories of Compensation." *Oxford Journal of Legal Studies* 9, no. 1 (1989): 56–75.

Harman, Elizabeth. "Can We Harm and Benefit in Creating?" *Philosophical Perspectives* 18, no. 1 (2004): 89–113.

98 I am grateful to Arthur Ripstein, Sophia Moreau, Brendan de Kenessey, Laura Valentini, Marina Moreno, audiences at the 2022 meetings of the Canadian Philosophical Association and the American Philosophical Association Eastern Division, participants at the 2022 Ludwig-Maximilians-Universität Colloquium in Political Philosophy, and five anonymous reviewers and editors at this journal.

Harrison, Caleb. "Supersession, Reparations, and Restitution." *Journal of Ethics and Social Philosophy* 19, no. 2 (2021): 148–67.
Helmreich, Jeffrey S. "The Apologetic Stance." *Philosophy and Public Affairs* 43, no. 2 (2015): 75–108.
Hindriks, Frank. "The Problem of Collective Harm: A Threshold Solution." *Erasmus Journal for Philosophy and Economics* 17, no. 1 (2024): 209–33.
Hurley, Paul, and Rivka Weinberg. "Whose Problem Is Non-Identity?" *Journal of Moral Philosophy* 12, no. 6 (2014): 699–730.
Johansson, Jens, and Olle Risberg. "A Simple Analysis of Harm." *Ergo* 9, no. 19 (2022): 509–36.
Kershnar, Stephen. "Are the Descendants of Slaves Owed Compensation for Slavery?" *Journal of Applied Philosophy* 16, no. 1 (1999): 95–101.
Kumar, Rahul. "Who Can Be Wronged?" *Philosophy and Public Affairs* 31, no. 2 (2003): 99–118.
———. "Why Reparations?" In *Philosophical Foundations of the Law of Torts*, edited by John Oberdiek. Oxford University Press, 2014.
Lambrecht, Felix. "On the Necessity of a Pluralist Theory of Reparations for Historical Injustice." *Philosophical Quarterly* 75, no. 3 (2025): 978–98.
———. "Pluralism, Structural Injustice, and Reparations for Historical Injustice: A Reply to Daniel Butt." *Ethical Theory and Moral Practice* 27, no. 2 (2024): 269–75.
———. "Reparative Justice, Historical Injustice, and the Nonidentity Problem." *Journal of Social Philosophy* (forthcoming). Published online ahead of print, July 18, 2024. https://doi.org/10.1111/josp.12583.
Lazar, Seth R. M. "Corrective Justice and the Possibility of Rectification." *Ethical Theory and Moral Practice* 11, no. 4 (2008): 355–68.
Lewis, David. "Causation." *Journal of Philosophy* 70, no. 17 (1973): 556–67.
Lu, Catherine. *Justice and Reconciliation in World Politics*. Cambridge University Press, 2017.
Mackie, John Leslie. *The Cement of the Universe: A Study of Causation*. Clarendon Press, 1974.
McKeown, Maeve. "Backward-Looking Reparations and Structural Injustice." *Contemporary Political Theory* 20, no. 4 (2021): 771–94.
Meyer, Lukas. "Intergenerational Justice." *Stanford Encyclopedia of Philosophy* (Summer 2021). https://plato.stanford.edu/archives/sum2021/entries/justice-intergenerational/.
———. "Past and Future: The Case for a Threshold Notion of Harm." In *Rights, Culture and the Law*, edited by Lukas H. Meyer, Stanley L. Paulson, and Thomas W. Pogge. Oxford University Press, 2003.
Meyer, Lukas, and Timothy Waligore. "Supersession and Compensation for

Historical Injustice." *Critical Review of International Social and Political Philosophy* (forthcoming). Published online ahead of print, March 17, 2024. https://doi.org/10.1080/13698230.2024.2309051.
Montero, Julio. "Colonialism and Rights Supersession: A Kant-Inspired Perspective." *Critical Review of International Social and Political Philosophy* 25, no. 3 (2022): 331–46.
Morris, Christopher W. "Existential Limits to the Rectification of Past Wrongs." *American Philosophical Quarterly* 21, no. 2 (1984): 175–82.
Nefsky, Julia. "Collective Harm and the Inefficacy Problem." *Philosophy Compass* 14, no. 4 (2019): e12587.
Nine, Cara. "Superseding Historic Injustice and Territorial Rights." *Critical Review of International Social and Political Philosophy* 11, no. 1 (2008): 79–87.
Nuti, Alasia. *Injustice and the Reproduction of History: Structural Inequalities, Gender and Redress*. Cambridge University Press, 2019.
Parfit, Derek. *Reasons and Persons*. Oxford University Press, 1984.
Paul, Laurie Ann, and Edward Jonathan Hall. *Causation: A User's Guide*. Oxford University Press, 2013.
Perry, Stephen. "The Moral Foundations of Tort Law." *Iowa Law Review* 77, no. 2 (1992): 449–514.
Pitcovski, Eli. "Explaining Harm." *Philosophical Studies* 180, no. 2 (2022): 509–27.
Purves, Duncan. "Harming as Making Worse Off." *Philosophical Studies* 176, no. 10 (2019): 2629–56.
Ripstein, Arthur. *Force and Freedom: Kant's Legal and Political Philosophy*. Harvard University Press, 2009.
———. *Private Wrongs*. Harvard University Press, 2016.
Sanderson, Douglas. "Redressing the Right Wrong: The Argument from Corrective Justice." *University of Toronto Law Journal* 62, no. 1 (2012): 93–132.
Sher, George. "Ancient Wrongs and Modern Rights." *Philosophy and Public Affairs* 10, no. 1 (1981): 3–17.
———. "Transgenerational Compensation." *Philosophy and Public Affairs* 33, no. 2 (2005): 181–200.
Shiffrin, Seana Valentine. "Reparations for US Slavery and Justice over Time." In *Harming Future Persons*, edited by David Wasserman and Melinda Roberts. Springer, 2009.
Simmons, A. John. "Historical Rights and Fair Shares." *Law and Philosophy* 14, no. 2 (1995): 149–84.
Slavny, Adam, and Tom Parr. "Harmless Discrimination." *Legal Theory* 21, no. 2 (2015): 100–114.
Song, Seunghyun. "Denial of Japan's Military Sexual Slavery and Responsibility for Epistemic Amends." *Social Epistemology* 35, no. 2 (2021): 160–72.

———. "Superseding Structural Linguistic Injustice? Language Revitalization and Historically-Sensitive Dignity-Based Claims." *Critical Review of International Social and Political Philosophy* 25, no. 3 (2022): 347–63.

Thompson, Janna. "Groups as Intergenerational Agents: Responsibility Through Time and Change." *Journal of Social Philosophy* 53, no. 1 (2022): 8–20.

———. "Historical Injustice and Reparation: Justifying Claims of Descendants." *Ethics* 112, no. 1 (2001): 114–35.

Waldron, Jeremy. "Superseding Historic Injustice." *Ethics* 103, no. 1 (1992): 4–28.

Weinrib, Ernest Joseph. *The Idea of Private Law*. Oxford University Press, 2012.

Wenar, Leif. "Reparations for the Future." *Journal of Social Philosophy* 37, no. 3 (2006): 396–405.

Wright, Richard W. "Causation in Tort Law." *California Law Review* 73, no. 6 (1985): 1735–828.

Journal of Ethics and Social Philosophy
vol. 30, no. 6 · october 2025
https://doi.org/10.26556/jesp.v30i6.4421

WHEN ENOUGH IS NOT ENOUGH

SATISFICING AND MORAL EFFICIENCY

Konstantin Weber

One way of distinguishing different moral theories is by their real-world demandingness. A theory is demanding (in the real world) to the degree that, as things actually are, adherence to it imposes burdens on us (taking into account both the prevalence and significance of burdens).[1] While burdens are usually understood as *welfare costs*, they may also be conceptualized as the *volitional effort* or *psychological difficulty* required in following the theory.[2] Although there is a substantive difference between these interpretations, I stay neutral between them in what is to come. Following a more demanding theory in the real world often makes one less well-off or requires more effort than following a less demanding one. According to an influential line of thought, this difference between theories is of great significance. Many people believe that there is a principled limit to the demandingness of morality. Usually, this is expressed by claiming that there is some absolute threshold of demandingness that morality does not exceed: while morality can and often is demanding, it cannot be overdemanding. This provides an argument against theories that cross this line: they fail to respect a central tenet of commonsense morality (so it is alleged). Attending to real-world demandingness allows us to demarcate a class of moral theories—namely, that of overdemanding moral theories—and reject them.

On the receiving end of this argument is, paradigmatically, classical utilitarianism.

1 Kyle York argues that the focus on real-world demandingness, as opposed to demandingness across the whole modal space, is misplaced, and he uses this to argue against cost-based rather than reasons-based interpretations of demandingness. Although I focus on real-world demandingness, I reach results very similar to York's. York's arguments, then, provide an additional reason to eschew absolute thresholds of burdensomeness. See section 3 below ("Moral Demandingness and Modal Demandingness").

2 Two further interpretations, which I only tangentially address, understand burdens as *cognitive requirements* needed to successfully apply the theory or as *restrictions of moral freedom* when following the theory—i.e., as the degree to which we can choose between permissible options.

> *Classical Utilitarianism*: An action is permissible if and only if its outcome is as good as possible (i.e., there is no available alternative action with a better outcome).

This view, it is often claimed, is almost at the top of the demandingness spectrum. As long as we believe that there is some relevant demandingness threshold, it should thus be abandoned. Given that this is one of the most prominent objections to classical utilitarianism, it is no surprise that utilitarians have tried a variety of responses. One of these is tactical retreat. Some utilitarians, conceding both that excessive demandingness is a problem and that utilitarianism does impose excessive demands, have opted to modify utilitarianism in a way that scales back the burdens it imposes. Instead of focusing on maximizing the good or always doing the best, as classical utilitarianism does, utilitarianism should be content with *satisficing*—that is, doing enough. Call this the *satisficing strategy*. Using this concessive maneuver, one could keep most of the general theory intact while escaping the objection from overdemandingness. This at least is the idea. In what follows, I investigate whether it can also become a reality.

It should be noted that while I focus on utilitarianism, the charge of being overly demanding and the corresponding pressure to limit the obligations imposed by one's moral theory are not restricted to utilitarianism.[3] For one, every theory that entails burden-insensitive positive obligations to benefit individuals or promote the good faces similar issues, from pluralist views that recognize a *pro tanto* obligation of beneficence to several versions of contractualism.[4] For another, even without positive obligations, a theory can be overly demanding by asking us to respect negative obligations. If I were lost in the mountains, and my only way of alerting the rescue service were to burn some of your belongings that I happen to be carrying, it would intuitively be overly demanding to require that I respect your right to not have your property damaged. As such, most of the arguments to follow can be, *mutatis mutandis*, applied to nonutilitarian theories as well.[5]

Before I go on, a caveat is in order. Although I speak of "the" objection from overdemandingness in this article, one can distinguish several substantively different complaints running under this rubric. I understand the objection from overdemandingness as targeted specifically at *moral* permissibility as

3 I thank an anonymous referee for inviting me to expand on this point.

4 Cf., for instance, Ashford, "The Demandingness of Scanlon's Contractualism."

5 I here assume that even on theories that do not allow for positive duties, there is a way to measure how much is morally at stake in a situation such that there can be a mismatch between the moral gain achieved by following the theory and the burdens imposed in doing so.

opposed to what we have most reason to do *overall*, or *all things considered*. To be clear, the arguments that follow do not hinge on this or on the specific way I understand the difference between moral and overall permissibility. They work even if one takes demandingness to be concerned with overall permissibility or prefers another account of how moral and overall permissibility relate.[6] I prefer this reading of the objection from overdemandingness because it preserves this objection as an independent complaint against utilitarianism that does not rely on a prior rejection of utilitarian, agent-neutral axiology (as the charge that utilitarianism does not allow for partiality to oneself, for instance, would).[7]

My own preferred view of the difference between moral and overall permissibility is *hypological*. Sometimes, if people act overall impermissibly without proper excuse, this makes it appropriate or fitting to have certain reactive attitudes towards them. These reactive attitudes (and their expression) constitute what we call moral blame. This moral blame, as usually understood, has characteristics that set it apart from other forms of criticism or reaction we can have in response to impermissible conduct, like thinking the other to be imprudent. The characteristic feature of moral as opposed to overall permissibility lies in exactly this connection with moral blameworthiness. This naturally leads to a specific interpretation of the objection from overdemandingness and, correspondingly, the satisficing strategy: they are concerned with the appropriateness of moral blame and praise.[8] The satisficing strategy attempts to formulate conditions of when overall impermissible behavior is, in addition, also morally wrong—i.e., when blame is appropriate—thereby capturing more of commonsense moral thought than classical utilitarianism.[9] This makes it possible to combine a satisficing version of utilitarianism on the moral level with a classically maximizing account of what we have most reason to do overall, which is why this understanding is particularly attractive for utilitarians.[10]

In this article, I evaluate extant versions of this satisficing strategy and propose a new version of it. After briefly dismissing in section 1 Michael Slote's *Satisficing Utilitarianism*, which is the best-known and most implausible version

6 For a focus on overall permissibility, cf. Case, "Rethinking Demandingness."

7 See also the discussion of Scheffler's view in section 4 below. I thank an anonymous referee for urging me to make this clearer.

8 This general idea is also pursued by McElwee, "Demandingness Objections in Ethics," 97–103; and Chappell, "Willpower Satisficing," 256–58.

9 An account of appropriate blame should not be confused with an account of permissible blame. That a person is morally blameworthy for some conduct does not imply that blaming this person is permissible. Indeed, utilitarians hold that it does not even give us any reason to do so.

10 Cf. Chappell, who argues for just this combination ("Willpower Satisficing").

of this strategy, I turn in section 2 to the more recent and much more promising interpretation of the satisficing strategy proposed by Richard Yetter Chappell. While Chappell's view is a big step forward, it does not solve all relevant problems because it misconstrues the objection from overdemandingness. In section 3, I argue *contra* Chappell that the objection from overdemandingness should not be read as claiming that there is some absolute limit to the burdens that morality imposes, for there is no such limit. Recognizing this, I propose in section 4 a new version of the satisficing strategy, *ratio satisficing utilitarianism*, that does without such an absolute limit and avoids the problems for Chappell's view. I then go on to develop and defend this view against some objections in sections 5 and 6.

1. Standard Satisficing

Slote's well-known version of the satisficing strategy is based on the introduction of a *utility threshold*. Once the outcome of an action passes this threshold, the action is automatically classified as permissible. Let us say that the outcome of an action that passes this threshold is *good enough*. Slote's view thus has the following general structure.

> *Utility Satisficing Utilitarianism*: An action is permissible if and only if its outcome is as good as possible *or* its outcome is good enough.[11]

In order to evaluate what this view implies, we need to say more about what 'good enough' amounts to. It is clear that to escape the charge of overdemandingness, the standard of good enough should lie noticeably below that of being the best. Proponents of utility satisficing utilitarianism trace the susceptibility of classical utilitarianism to overdemandingness worries back to its maximization requirement. Maximization, they claim, is often excessively burdensome: doing the best is far too hard. In retreating to the lower bar of doing enough good, several views about what is good enough are possible, most notably an absolute interpretation, according to which an outcome is good enough if it contains a specified absolute amount of utility (either on the whole or for each individual), and a relative interpretation, according to which an outcome is good enough if it is sufficiently close to the best outcome available. While each specification of this threshold introduces specific problems of calibrating between being overly and insufficiently demanding, utility satisficing

11 Cf. Slote, "Satisficing Consequentialism" and *Common-Sense Morality and Consequentialism*. The first clause is needed to handle cases in which no outcome is good enough. Whether such cases are possible depends on how we spell out the notion of a good enough outcome.

utilitarianism fails because of a more fundamental difficulty that is independent of how we understand 'good enough.'[12]

This difficulty is that utility satisficing utilitarianism is structurally inadequate to answer the objection from overdemandingness. Utility satisficing utilitarianism attempts to scale back the burdens imposed by utilitarianism by restricting the good that has to be brought about. But the goodness of an outcome is not always proportional to how burdensome it is to bring this outcome about. Because of this *problem of incongruity*, utility satisficing utilitarianism grossly misclassifies the status of many actions. Utility satisficing utilitarianism leads to over- or underdemandingness independently of where the utility threshold is set. If it is low, then it is often permissible not to perform barely burdensome actions with amazing outcomes. If it is high, then it is often impermissible to refrain from highly burdensome actions with outcomes that are barely better than those of much less burdensome alternatives.

This misclassification is tied to the fact that utility satisficing utilitarianism allows egregious *moral inefficiency*. Since all actions with outcomes above the utility threshold are permissible, it is possible to choose some such action even though there is a better, less burdensome alternative. But such gratuitous prevention or foregoing of utility is reprehensible. We would therefore expect that utilitarian theories are committed to the following:

> *Gratuitous Goodness*: An action is impermissible if there is an equally or less burdensome alternative with a better outcome.

I argue later that gratuitous goodness is in fact ambiguous and only plausible as long as we understand "equally or less burdensome" in a particular, narrow way. What is true, in any case, is that egregious inefficiency must be off the table. Noticing these problems, it should be clear that utility satisficing utilitarianism is entirely unable to provide an answer to the objection from overdemandingness. As a response to the objection from overdemandingness, it is without theoretical promise.[13]

12 I discuss the problems arising from different specifications of the relevant threshold in more detail in Weber, "Schwellenwertutilitarismus."

13 There are two other projects for which utility satisficing utilitarianism might nevertheless be useful. First, utility satisficing utilitarianism could have some virtues as an indirect decision procedure intended to reduce cognitive demands on deliberating agents (though I doubt utility satisficing utilitarianism would be the best candidate for this). Second, utility satisficing utilitarianism or something like it might be appealed to as an account of a specific moral subdomain like justice. Utility satisficing utilitarianism might, for instance, spell out what each individual has a basic moral claim to. Both of these interpretations of utility satisficing utilitarianism are quite different from the one I give in the text and play no role in this article.

2. Sophisticated Satisficing

The failure of utility satisficing utilitarianism should not sour us on the satisficing strategy in its entirety. Richard Yetter Chappell proposes another version that holds far more promise. Chappell argues that we should introduce a threshold not of utility but of burdensomeness. The problem with classical utilitarianism is not that it is maximizing but that it maximizes *without constraints*. His view, which works as follows, is intended to remedy this.

> *Burden Satisficing Utilitarianism*: An action is permissible iff every alternative with a better outcome is *both* (1) excessively burdensome *and* (2) more burdensome.[14]

According to Chappell, each action has some level of absolute burdensomeness that can be specified. And there is some threshold along this scale such that actions above it are *excessively burdensome*. This leads Chappell to understand the objection from overdemandingness, as directed toward classical utilitarianism, as follows: it complains that classical utilitarianism sometimes classifies only excessively burdensome actions as permissible. I argue later that this understanding is mistaken. But first, we should see how Chappell answers it. Chappell's view makes use of two components. First, Chappell proposes a *burden threshold* instead of Slote's utility threshold (this is borne out by condition 1): it is impossible for an overdemanding action to be the sole permissible option (thus being obligatory) since there is always a permissible option below the burden threshold. In essence, 1 gives us a sufficient condition for any action to count as morally permissible—namely, having (one of) the best outcome(s) of all actions whose burdensomeness falls below the threshold. Burden satisficing utilitarianism thus ensures that there will always be at least one permissible action that does not impose high absolute burdens (if there is some such action).

A second difference to Slote is that Chappell does not rest content with just the threshold. It is not true, Chappell argues, that every action above the threshold with a better outcome than the best below the threshold is permissible. Instead, there is an additional, necessary condition for permissibility: there must be no better alternative that is not also more burdensome (as condition 2 states). This condition ensures that, for all actions that share the same level of burdensomeness, only those with the best outcomes are permissible.[15] This

14 Cf. Chappell, "Willpower Satisficing," 252. I adjust Chappell's formulation to make it easier to read.

15 More precisely, an action at a certain level of burdensomeness is permissible only if there is no alternative with a better outcome that is equally or less burdensome.

leads to the existence of *supererogation holes*: cases where actions that are better than the minimum are nevertheless impermissible.[16] The cases of egregious moral inefficiency provide prime examples of this.

Chappell's view is superior to Slote's in every respect. First, Chappell's first condition directly ties the permissibility of an action to the level of burdensomeness of that action. This avoids the problem of incongruity. Second, Chappell avoids the abhorrent implications of Slote's view by directly incorporating gratuitous goodness—which is identical to Chappell's second condition. Chappell's view is a major improvement over utility satisficing utilitarianism.

Not so fast. Although burden satisficing utilitarianism is a major step forward, it is not beyond reproach. In two recent publications, Joe Slater raises two problems for Chappell's view.[17] Slater's first objection consists in a dilemma faced by gratuitous goodness. The gratuitous goodness principle, Slater complains, allows for two different readings depending on how we understand 'equally burdensome'—namely, either narrowly or broadly, each of which has problematic consequences. If we adopt the narrow understanding, then the scope of the gratuitous goodness principle is overly restricted, and it fails to apply in all cases in which we would expect it to apply. If we adopt the broad understanding, however, then the gratuitous goodness principle is too strict and applies in too many instances. There is, then, no way to correctly calibrate the principle.

Let us look at this objection in a little more detail.[18] Imagine that, moral exemplar that you are, you decide to donate all your savings to an effective charity. Your donation is so substantial, we can suppose, that it goes far beyond the relevant burden threshold such that it is not obligatory for you to donate this sum. There are two different charities under consideration. One charity will use the money to combat world hunger, while the other will fund important medical research on some illness. Let us assume that, as you know, both charities will do quite similar amounts of good, though the second charity will achieve a little less. Your mother, whom you admired, was a medical researcher driven by her dream to eradicate the very same illness; you identify with the vision and approach of this charity and would prefer donating to the second charity because of this. How ought you spend the money?

The answer of burden satisficing utilitarianism depends on whether the gratuitous goodness principle applies (since donating to either charity would be

16 For this expression, see Wessels, "Beyond the Call of Duty," 90. For a treatment of this feature of supererogation, see also Horton, "The All or Nothing Problem."

17 Slater, "Satisficing Consequentialism Still Doesn't Satisfy" and "Satisficers *Still* Get Away with Murder!"

18 See Slater, "Satisficing Consequentialism Still Doesn't Satisfy," 114–16.

overly burdensome), i.e., on whether both donations are *equally* burdensome. Given how the case is stipulated, this is not entirely clear: while both donations are equally substantial, you have a preference for one of them, after all. According to a first possible interpretation, such minor differences (e.g., slight preferences, different emotional connections, or insubstantial additional exertions of willpower) do not influence an action's relevant level of burdensomeness. On this *broad understanding* of burdensomeness, levels of burdensomeness are individuated in a rather coarse-grained way, and the requirements for gratuitous goodness to apply are low. So too in the case at hand: given that there is at most a slight difference in the things relevant for burdensomeness, it would be impermissible to follow your preference and donate to the second charity closer to your heart.

This is not, however, the answer that strikes many as correct. For many, it seems clear that it should be permissible to choose the second charity, in line with your preferences. Slater, for instance, comments, "Many of us would recoil at the prospect of condemning anyone who gives a substantial amount of their income to a non-perfectly-efficient charity—perhaps even the second best charity.... If someone did give ... to the second best charity, it would seem extremely inappropriate to condemn them for doing so."[19] The broad understanding of gratuitous goodness is thus much too strict: it robs us of some moral leeway in going beyond the call of duty that, intuitively, we seem to have.

This suggests that we should give up the broad understanding and adopt a *narrow understanding*, according to which levels of burdensomeness are individuated in a fine-grained way, and the requirements for gratuitous goodness to apply are high. This would lead to the correct result in the case at hand: both donations would be permissible. But it would lead to even more implausible verdicts in other cases. It would, for instance, be permissible to save two rather than three people from a burning building if saving anyone is above the burden threshold, and we do not particularly like the third guy. Given our slight preference for not bothering with him, saving two and saving three would not be equally burdensome, and we would, in saving the two, maximize utility on that level of burdensomeness. This obviously cannot be right. But then, the gratuitous goodness principle is either false or too weak to do the work it is intended to do, which leads to a dilemma for burden satisficing utilitarianism.

We can now turn to Slater's second objection (which is already anticipated though not adequately addressed by Chappell).[20] It concerns the burden threshold. For burden satisficing utilitarianism struggles with cases that have the

19 Slater, "Satisficing Consequentialism Still Doesn't Satisfy," 114.

20 Cf. Chappell, "Willpower Satisficing," 255; and Slater, "Satisficers *Still* Get Away with Murder!" 1365–68.

following structure: there is an action that is (a) very slightly above the burden threshold and (b) generates significantly more utility than any action below the burden threshold. Consider, for instance, the following macabre variation of a trolley case. An out-of-control trolley is rushing towards a group of innocents. There are two bridges above the track, with a heavy man on one and a large sandbag on the other, both of which would stop the trolley if pushed onto the track. You have unfortunately been ensnared in a bear trap. You could free yourself and rush onto the bridge, but this would cost you one of your legs and would make all movement considerably painful. The bridge with the man is closer to you than the one with the sandbag on it, though you could reach both in time. You now have three options: (1) stay put and let the innocents die; (2) sacrifice your leg, endure the pain of hobbling toward the closer bridge, and push the man, saving the innocents; or (3) sacrifice your leg, endure the pain of hobbling toward the more distant bridge, and push the sandbag, saving the innocents. If the distance between the bridges is just large enough to make option 2 and option 3 unequally burdensome, then burden satisficing utilitarianism would classify all three options as permissible. But given the comparatively small difference in burdensomeness and the comparatively large difference in utility between 2 and 3, this seems counterintuitive: 2 should be impermissible![21] You would, it seems, be open to substantial moral reproach should you sacrifice your leg but be unwilling to endure a few extra moments of pain to avoid sacrificing a life.

To these two flexibility-based objections, we can add another. Burden satisficing utilitarianism also has the implication that significant additional burdens have to be borne in order to secure tiny improvements in utility if we are below the burden threshold.[22] Imagine that you face the choice between

21 Chappell is aware of this and writes that the level of maximum burdensomeness "plausibly ought to be context-sensitive, e.g. to ensure (1) that it's never permissible to do just a little good when a huge amount of good could be achieved by an only *slightly* more [burdensome] action" ("Willpower Satisficing," 255). However, Chappell fails to provide an account of how exactly this should be done. (In fact, the way he spells out his account seems to preclude this context sensitivity.) It seems that by sticking to the threshold conception, Chappell's options in addressing this problem are very similar to the options of a proponent of utility satisficing utilitarianism in addressing the problems with that theory's threshold. Furthermore, adopting strong moral efficiency is incompatible with the existence of absolute limits of imposable burdens (cf. Wessels, "Beyond the Call of Duty," 93–95), something that Chappell apparently wishes to uphold.

22 Douglas Portmore raises a similar complaint against *classical utilitarianism*—namely, that it implies "agents should sacrifice, not only their disposable income but even their own lives and the lives of those whom they love most whenever doing so will produce the most aggregate utility, and thus, even when the net gain would be as small as one utile" (where a "utile" is the smallest possible unit of utility) (*Commonsense Consequentialism*, 4). Note that Portmore's main grievance lies not in the fact that utilitarianism can impose such

two alternatives: the first would not impose any burdens on you and secure a great amount of goodness; the second would, however, be noticeably burdensome—namely, such that it would be just shy of being excessively burdensome in Chappell's framework. Still, the second would secure just a tiny bit more utility (the burdens imposed on you already accounted for). Burden satisficing utilitarianism and classical utilitarianism would imply that because of this, it would be impermissible to choose the first action. But that seems implausible: for the first action is only very slightly worse but far less burdensome.

In all three cases, burden satisficing utilitarianism appears to be insufficiently flexible to allow for a proper picture of moral permissibility. What should we make of this?

3. Against the Burden Threshold

The three objections just considered raise valid complaints: burden satisficing utilitarianism really is insufficiently flexible. However, just leaving it at that, though perhaps adding some epicycles, fails to target the core of the matter. I believe that the source of these problems concerns the very heart of Chappell's view: the absolute burden threshold. Chappell, in line with many others, seems to adopt an absolute interpretation of the objection from demandingness: on this understanding, actions below some level of burdensomeness are never problematic with respect to demandingness while actions above some level of burdensomeness are always problematic with respect to demandingness. Both claims are, I think, mistaken: there are no absolute limits of morally acceptable burdensomeness. Like Slote's utility threshold, Chappell's burden threshold needs to be abandoned. After making the case for this claim, I go on to look at its implications for the satisficing strategy and propose a better version of it: instead of restricting the demands for either utility or burdens, we need to look at the proper balancing of both.

Adopting an absolute burden threshold involves subscribing to two claims. On the one hand, one accepts that below this threshold, there is no demandingness-related objection to moral obligatoriness. If burdens are low, then we really ought to do the best. On the other hand, one accepts that there is an absolute limit to morality's demands. Actions that, in absolute terms, reach a sufficient level of burdensomeness can never be obligatory. Both claims are, I believe, mistaken.

---

burdens but in the fact that it imposes them to secure only trivial improvements: "I find it incredible to suppose that such miniscule gains in aggregate utility could be sufficient to make it reasonable to perform such acts" (4).

On the one hand, it is sometimes permissible not to perform very undemanding actions with (slightly) better outcomes than available undemanding alternatives. Consider the following.

> *Flowers*: Due to a billing error, I happen to have an extra bouquet of flowers in my possession, which I have no real use for. I could bring them over to my elderly neighbor, to whom they would bring some (though not great) enjoyment. However, since the weather is bad, and my chair is comfortable, doing so would be slightly burdensome for me.

Given that gifting the bouquet is surely not excessively burdensome and better than the alternative, Chappell's view implies that it would be impermissible not to do so. However, it seems to me that this can be criticized on grounds that are very similar to the worries voiced by the demandingness objection. My failure is much too trivial and insignificant to merit moral reproach, and subjecting it to a stringent moral demand would be excessive. It does seem overly demanding to require us to endure some burdens for, in comparison, quite trivial improvements. Such a demand is excessive not because of the absolute size of the burden it imposes but because of the insignificance of the good it secures. This is especially true if we understand the objective of the satisficing strategy as supplying part of a condition of moral blameworthiness (as Chappell also does). It does not seem appropriate to react with moral blame toward my decision not to gift the flowers, and this cannot plausibly be explained by some excusing or exculpating condition. Thus, Chappell's view is still too demanding in this way: it requires actions that make only trivial improvements.[23]

One might be tempted to reply that instead of not being blameworthy at all, I am blameworthy only to a minuscule degree. It is often claimed to be difficult to distinguish between cases where a normatively relevant factor is not present at all and cases where it is present to a very small degree. I have two things to say. First, I do not believe that it really is difficult to distinguish between no presence and minor presence when it comes to moral blameworthiness. One of the characteristic features of moral blame is its seriousness. Moral blame stands out among all reactive attitudes as particularly weighty: we cannot (and usually do not) "shrug off" appropriate moral blame in the same way we can (and often do) shrug off a charge of impoliteness or bad taste. Blame "carries a

23 Note that this concerns also actions where the stakes are not trivial for everyone. If I could perform some burdensome though not excessively burdensome action that would cause someone else a little more good than I would myself sacrifice, it would be implausible to claim that I am morally obligated to perform it (if we take demandingness considerations seriously).

characteristic depth, force, or sting."[24] Because of this, we are generally speaking very receptive as to when we are fitting targets of it. If so, cases in which appropriate moral blame is barely recognizable become dubious: once we enter into the realm of moral blame, we should and usually do take it seriously. This does not mean of course that there is no distinction between strong and weak moral blame (or blameworthiness), only that even being morally blameworthy to a minor degree is still quite noticeable.[25] Second, even if we allow for the possibility of moral blame sometimes being hard to spot, it seems to me that the not-blameworthy-at-all side has an intuitive advantage. It would not seem appropriate for me to feel bad at all for not bringing over the flowers; it could clearly not be reasonably demanded of me; and my neighbor would be making a mistake were they to resent me for it. We should acknowledge that demands can be objectionable not because of the seriousness of the imposed burden but because of the insignificance of the achieved moral good.

That such actions are not usually taken to fall into the sphere of obligation is also evidenced by the fact that, for instance, gestures of kindness, small favors, or making an effort at being extra friendly are classified as *supererogatory* by commonsense morality. Ordinary moral life is, with regard to such actions, characterized by a form of moral freedom: we are free to perform such actions but will not generally be blameworthy if we do not put in the extra effort. Quite the contrary, our performance of such actions ordinarily merits moral praise. If I were to gift the bouquet to my neighbor, it would be fitting for the neighbor to feel gratitude in a way that need not be fitting if I were simply discharging a clear moral requirement. If one aims at avoiding the demandingness objection in general, one should also aim at avoiding a requirement to secure goods that are, in comparison to the burdens incurred, trivial. Doing so requires, however, the rejection of a burden threshold below which the burdens imposed by actions cease to be of moral interest. Chappell's view, being built around such a threshold, cannot be fully adequate.

On the other hand, it is sometimes impermissible not to perform very demanding actions.[26] Consider the following.

> *Train*: Through no fault of my own, I happen to be fixed onto some train tracks. An express train, fully occupied by many innocents, is speeding

24 Hieronymi, "The Force and Fairness of Blame," 117.

25 I thank an anonymous referee for pressing me on this point.

26 Note that this is already implied by the strong version of moral efficiency I advocate: there could be a sequence of ever more burdensome but still efficient actions, where, given that we ought to perform one action, we also ought to perform its "successor," independently of how absolutely burdensome this successor is.

> towards me, threatening to crush me to death. The only way to avoid being crushed is to engage a mechanism that would blow up the whole train, killing most and maiming some of the passengers. Though overcome by immense fear and trepidation, it is, with considerable effort, possible for me not to engage the mechanism.

Surely, not engaging the mechanism is excessively burdensome in absolute terms. Chappell's view would classify this action as permissible but not obligatory. Nevertheless, it is plausible that I am morally obligated not to engage the mechanism, and this appears to be the verdict that commonsense morality gives as well. The same can be said about other cases in which our refusal to shoulder a certain burden leads to a moral catastrophe: most moral theories plausibly imply that it is impermissible to let such a catastrophe come to pass. Moreover, it seems to me that commonsense morality requires undertaking similar burdens even if the stakes are substantially lower. If some stranger and I happen to be in need of some lifesaving drug, and this drug is in the stranger's possession, commonsense morality pronounces it impermissible to kill this stranger, even if this is the only way to save my life. In general, then, even according to commonsense morality, there does not appear to be any absolute limit to morality's demands.[27] Since Chappell builds a burden ceiling into his theory, it classifies too many actions as overdemanding. Chappell's reading of the objection from overdemandingness is much too rigid, and his resulting view imposes too little burdens as a result. Again, burden satisficing utilitarianism does not appear to be the best version of the satisficing strategy available.

Of course, critics might jump in at this point. There is, they might claim, an absolute burden ceiling because we have a fundamental egoistic prerogative to secure a sufficiently decent life for us that enjoys lexical priority over other normative considerations. It would then always be permissible to prioritize our interests over those of others given that our interests are sufficiently affected. But such a prerogative is not something that, as the original objection from overdemandingness, can be the premise of a convincing argument against allegedly overly demanding theories: it lacks robust grounding in common sense and is itself a seriously controversial claim in need of justification. If our aim is to move classical utilitarianism—or any other theory thought to be overly demanding—closer to commonsense morality by scaling back the burdens it imposes, we need not address views that depart from commonsense morality in imposing even stricter limits. And if we wish to formulate a

27 Additional support for this can be found in York's modal conception of demandingness in "Moral Demandingness and Modal Demandingness."

convincing objection to classical utilitarianism, then we should not appeal to principles that themselves have feet of clay.

4. Efficient Satisficing

We can develop a better version of the satisficing strategy and a better interpretation of the demandingness objection by heeding the lessons we have learned from discussing utility satisficing utilitarianism and burden satisficing utilitarianism. What we need is a view that is thoroughly committed to moral efficiency without positing arbitrary absolute limits of burdensomeness. An action is morally efficient in the sense I am interested in if and only if in performing the action, we get a sufficient amount of additional goodness for a comparatively smaller increase in burdensomeness. What this means is that we should see to it that the burdens imposed by some action and the goodness achieved by it stand in the right relation. This leads us directly to the following proposal.

> *Ratio Satisficing Utilitarianism*: An action is permissible if and only if there is (1) no alternative that is equally or less burdensome but has a better outcome *and* (2) no more burdensome alternative with a better outcome that has a sufficient burden-utility ratio relative to it.[28]

The *burden-utility ratio* of an action relative to another more burdensome action with a better outcome tells us how much additional good we get in relation to the additional burdens we undertake:

28 Slater considers and rejects a modification of Chappell's view that shares some similarity with ratio satisficing utilitarianism:

> Restricted Effort-Satisficing (RES): An act ϕ is permissible iff:
>
> 1. ϕ brings about at least as much good as could be achieved by expending X effort.
>
> AND
>
> 2. If the agent expends more than X effort, there must not be an option ψ requiring a similar amount of effort that brings about significantly more good than ϕ. . . .
>
> Another option for interpreting 'significantly more' would be as a function of the marginal utility (or marginal improvement to the consequences, for non-utilitarian consequentialists) *and* the marginal effort. We might, for instance, divide the marginal utility of an alternative action (in hedons) by the marginal effort costs (in effort points), and if the result is sufficiently high, we could regard this action as doing *significantly* more good. ("Satisficers *Still* Get Away with Murder!" 1370–72)

However, this view still incorporates an absolute burden threshold.

$$\textsc{bur}(a, b) = \frac{goodness(a) - goodness(b)}{burdens(a) - burdens(b)},$$

with the numerator denoting the difference in goodness between the (outcomes of the) actions and the denominator denoting the difference in absolute levels of burdensomeness between the actions.[29] An alternative has a sufficient burden-utility ratio relative to the action under consideration iff $\textsc{bur}(a, b)$ passes a certain threshold. Like utility satisficing utilitarianism and burden satisficing utilitarianism, then, ratio satisficing utilitarianism makes use of a threshold. Unlike utility satisficing utilitarianism and burden satisficing utilitarianism, however, this is a threshold not of utility or burdensomeness alone but of their relation. The fraction giving us this relation is defined only if the actions under consideration differ with respect to their burdensomeness, and we can restrict it to cases where this is so. In all other cases—i.e., in the comparison of equally burdensome actions—condition 1 kicks in.[30] Ratio satisficing utilitarianism is thus well defined for all cases.

Ratio satisficing utilitarianism is similar to some views found in the literature. It is structurally identical to Samuel Scheffler's *hybrid theory*, which combines a concern for the impartial good with granting self-standing importance to the agent's personal point of view. In spelling out such a theory, Scheffler makes the following proposal:

> I believe that a plausible agent-centred prerogative would allow each agent to assign a certain proportionately greater weight to his own interests than to the interests of other people. It would then allow the agent to promote the non-optimal outcome of his choosing, provided only that the degree of its inferiority to each of the superior outcomes he could instead promote in no case exceeded, by more than the specified proportion, the degree of sacrifice necessary for him to promote the superior outcome. If all of the non-optimal outcomes available to the

29 Since classical utilitarianism and its cognates typically assume that goodness is an extensive quantity measurable on a ratio scale, for convenience's sake, I also work under this assumption. For my purposes here, however, it suffices to have interval scale measurability.

30 Condition 1 is identical to the narrow understanding of Gratuitous Goodness. As I argue, this principle poses problems for burden satisficing utilitarianism—namely, that it is far too permissive once we are above the burden threshold. But the same problem does not arise here. This is because cases where burdens are not exactly equal are covered by condition 2. Of course, some might still think that choices where 2 does not apply should be classified as equally permissible because they are morally trivial. While I do not agree with this, those sympathetic to this claim can replace 1 with a corresponding condition. Note that some condition is needed to cover cases where 2 is not defined.

> agent were ruled out on these grounds, then and only then would he be required to promote the best overall outcome.[31]

Like with ratio satisficing utilitarianism, Scheffler proposes to attend to the relation between the burdensomeness of an action and the utility it generates (with burdens interpreted in terms of welfare costs): we are required to take on additional burdens only if this secures a significantly greater gain in additional goodness. We want "goodness bang" for our "burdens buck." In substance and especially in its justification, however, there are noticeable differences. Scheffler develops his view from the guiding idea that the personal point of view is normatively independent from the utilitarian's "point of view of the universe" and has self-standing normative importance: an agent is allowed to weigh their own interests more than the interests of others. Because of this, Scheffler's proposal has general implications for the objection from overdemandingness and the satisficing strategy. If Scheffler were right, then the failure the objection from overdemandingness attributes to classical utilitarianism would be a mistaken underlying axiology or view about what we have most reason to do.

Understanding the objection from overdemandingness in this way raises three related problems (some of which have previously come up already). First, it undermines the status of this objection as a self-standing complaint that is independent of quarrels about moral and nonmoral reasons (for actions) or axiology.[32] Instead, it relies on a prior rejection of utilitarian axiology—which would of course need to be substantiated by some independent argument to not be simply question begging.[33] Second, because of this, taking a concessive approach with regard to this interpretation of the objection from overdemandingness is quite unappealing for utilitarians. I said at the outset that the objection from overdemandingness, as I want to understand it here, concerns specifically moral permissibility, leaving room for a classic utilitarian answer to the question of what we have most reason to do. But once we acknowledge that the objection targets the underlying utilitarian axiology, this view becomes untenable. If we were to bury the utilitarian commitment to a thoroughly agent-neutral axiology to account for certain moral phenomena, we could not comfortably resurrect it to account for what we have most reason

31 Scheffler, *The Rejection of Consequentialism*, 20.

32 A similar point is also made by Chappell, "Willpower Satisficing," 254.

33 For this line of criticism of the objection from overdemandingness, see again Sobel, "The Impotence of the Demandingess Objection." One possible reply is that the intuition pumped by cases characterized as overdemanding just is part of an argument against agent neutrality: they just illustrate one specific downside of it. I do not pursue further the question of how convincing the reply is here.

to do, all things considered. Third, in making room for the independent point of view of the agent, which amounts to counting their welfare differently than the welfare of others, Scheffler is forced to understand burdens as welfare costs, robbing the proposal of some welcome neutrality with regard to this question.

Ratio satisficing utilitarianism, on the other hand, is not wedded to the idea that the agent's welfare has some independent normative significance. It is driven not by general axiological principle but by the specific intuitions related to the demandingness of theories, however this is understood in detail in the realm of morality. Although this must be spelled out a lot more, ratio satisficing utilitarianism is compatible with taking the objection from overdemandingness to be independent from disputes about axiology and therefore allows for a fully utilitarian account of what we have most reason to do, all things considered. At the same time, because of ratio satisficing utilitarianism's introduction of a moral permissibility-relevant threshold, it is still a riff on the idea of satisficing—i.e., doing enough without doing the best.

It is also clearly distinct from the previous versions of the satisficing strategy. It not only incorporates the gratuitous goodness principle and sidesteps the issues of utility satisficing utilitarianism but also avoids the problems with moral inefficiency that burden satisficing utilitarianism struggles with. For one, ratio satisficing utilitarianism gives us the correct verdicts in Flowers and Train. Since we may suppose that gifting the flowers to my neighbor is slightly more burdensome and only a little bit better than not giving the flowers to my neighbor, it plausibly falls below the efficiency threshold and is thus not obligatory (even though it would be better). And since blowing up the train would be disastrous, the substantial additional burden that condemning myself to certain doom brings with it appears to be gained back many times over in goodness, bringing it above the threshold and thus rendering it obligatory.

Moreover, ratio satisficing utilitarianism also avoids the three general problems that burden satisficing utilitarianism faces. First, like with burden satisficing utilitarianism, the question of how to individuate levels of burdensomeness arises. Ratio satisficing utilitarianism allows, other than burden satisficing utilitarianism, for an easy solution: we should adopt the fine-grained principle for individuation. This makes the gratuitous goodness principle clearly true without leading to the unacceptable verdicts of burden satisficing utilitarianism. If there is only a slight difference in burdensomeness with a substantial difference in goodness, as in the trolley example above, then the respective return on investment is sufficiently high to mandate the alternative with the better outcome. Second, it allows for leeway by, for instance, holding that the choice of the second charity is permissible because the additional goodness gained from the other need not be sufficient to justify the additional though very minor burdens. Third, it does not force us to exert ourselves up to a certain point just in order to achieve trivial improvements. All told, ratio satisficing utilitarianism shines exactly where burden satisficing utilitarianism stumbles.

5. Discounting Ratio Satisficing Utilitarianism?

In this and the next section I address two possible objections to ratio satisficing utilitarianism. One thing I do not commit to is any specific value for what the burden-utility ratio should be. Even beyond the general challenge of providing cardinal measurements for utility and for burdensomeness, deciding on a value for this requires careful consideration of different cases in combination with a more substantiated conception of burdensomeness than I can develop here. Still, I want to argue for a structural claim about this value: it is constant and thus insensitive to where on the absolute scale of burdensomeness the actions that we are comparing lie. This argument is necessary because some might initially be drawn to the opposite claim. To see why, it is helpful to look at cases with a peculiar structure: we start with a baseline action that imposes no burdens at all. Call it a_1. We then add to the alternatives an action a_2 that, relative to a_1, has greater absolute burdensomeness and secures an outcome better by some absolute amount, with $\text{BUR}(a_1, a_{1+1})$ being sufficiently high to make a_1 impermissible. We can now iterate this procedure by introducing a further alternative a_3, which differs from a_2 by the same absolute amounts of goodness and burdensomeness as a_2 differs from a_1, and so on. With this, we can reach obligations that impose arbitrarily high burdens.[34]

To illustrate, take the following example.[35] You can donate different sums of money, and we are trying to figure out how much you are morally obligated to donate. Let us say that donating €50 to save one life is worth it. Your donations have the following effects:

Donation in €	Number of Lives Saved
0	0
50	1
100	2
150	3
…	…
1,000	20
1,050	21

34 For a structurally similar case, cf. Wessels, "Beyond the Call of Duty," 93–95.

35 This example is adapted from Wessels, "Beyond the Call of Duty," 94.

It seems that if the first donation is worth it (i.e., if we ought to move from €0 to €50), then we also are obligated to take every further step if we adopt a constant threshold for the burden-utility ratio. Depending on how much money you have maximally available, it seems that we could reach the point where you are obligated to give it all up.

Now, given that my arguments against the existence of an absolute burden threshold are sound, *that* we can reach burdens of this level should in itself not be objectionable. But we might have a more subtle worry about ratio satisficing utilitarianism that cases like this bring into the light—namely, *how fast* we reach this level of burdensomeness. In keeping the threshold for a sufficient burden-utility ratio constant, we treat any step of this iterative procedure exactly like the first. Even if in absolute terms, we are already dealing with a highly burdensome action, the improvements that we are required to bring about once the next alternative has been introduced is the same as the improvements that the first alternative we have introduced has brought. And this might seem implausible: Should we not be sensitive to the fact that a few steps into the procedure, we demand things of someone who already is under enormously burdensome obligations? Should we not give more leeway to those who already shoulder large burdens than we give to those who have yet to strain their backs?

Many would, it seems, answer positively. Ratio satisficing utilitarianism would then improperly balance burdens and marginal goodness. If we share this impression, there is a remedy in reach. Following Ulla Wessels, we might argue like this:

> Supererogation boils down to granting the agent, as an agent, a *discount* in matters of morality. Yet we need not stop at a discount. We could also give the agent a *bulk* discount, an additional discount for cases in which she is already burdened. Hence, we could ascribe to the agent for the *same* moral yield fewer additional obligations the worse off she is or the more she has already done for morality.[36]

Wessels's idea is that the required burden-utility ratio should be higher the greater the absolute levels of burdensomeness of the actions compared are: if the actions compared are not very burdensome in absolute terms, the sufficiency threshold for the return on investment is lower, while it is higher if the actions compared are very burdensome in absolute terms. This still avoids positing an absolute limit to the burdensomeness of an obligatory action but slows down the acceleration of burdens: the higher on the burdensomeness spectrum we go, the higher the rewards must be to obligate us to move further.

36 Wessels, "Beyond the Call of Duty," 95.

Incorporating this idea into ratio satisficing utilitarianism is possible, but it makes the whole thing a little cumbersome. One way to do it is to introduce an exponential function $f\colon x \mapsto c^x$ and apply it to the levels of burdensomeness. We would then calculate the burden-utility ratio like this:

$$\textsc{bur}(a, b) = \frac{goodness(a) - goodness(b)}{f(burdens(a)) - f(burdens(b))}.$$

Since f grows exponentially, the difference in the f-values get larger the higher the initial values of absolute burdensomeness are, lowering the value of the burden-utility ratio that we get in the end.

I do not see this modification as really necessary. The impression that a flexible sufficiency threshold is needed stems, I think, from consideration of "real-world" cases like the donation example above. And in these cases, it appears immensely plausible that there is a difference between going from donating nothing to donating €50 and going from donating €1,000 to donating €1,050. That is certainly true. But it is also perfectly compatible with a constant threshold. The reason is that the amount of money we spend does not have a fixed exchange rate to individual utility. Instead, money has diminishing marginal utility. If we understand burdens as welfare costs, then this already explains why even with a constant threshold for how much welfare we must be prepared to sacrifice for a certain moral gain, the first increase in the amount donated being obligatory is easier to justify than the five-hundredth increase being obligatory. Due to the diminishing marginal individual utility of money, the welfare costs in making this improvement go up. And insofar as this holds not only for money but for most of the resources we can use to promote utility (like time, material goods, hard work, etc.), similar considerations apply to those. Because we do not directly exchange our own welfare for general utility but rather exchange it indirectly via welfare-giving resources, and because these resources often have diminishing marginal individual utility, a constant sufficiency threshold for additional welfare sacrifices leads to a flexible rate on the level of resources. Unless personal utility itself exhibits some form of discontinuity or dampening, I see no reason why we should allow for flexibility on the level of individual and general utility.

But what if we were to understand burdens differently—namely, as needed effort or willpower? I do not think this would change much. Let us again look at a (very stylized) real-world example. Instead of monetary donations, we can now think of the relevant burden as hours of strenuous labor (in the sense that doing this labor requires willpower). Now, either it is the case that once we have already put twenty hours in, putting in another hour requires more willpower than the first took—in which case the situation is exactly the same as above because it is not the sufficiency threshold that changes but the amounts inputted—or putting in the twenty-first hour takes exactly as much willpower, and we still have this amount of willpower left. Then I fail to see what counts in favor of treating this differently from the first: the moral yield is, after all, well worth it. I therefore see no reason to modify ratio satisficing utilitarianism along the proposed lines.

6. Trouble at the Border?

Views that posit some sort of threshold, like the satisficing views I consider here, are often subject to problems when we consider cases in which stuff happens close to this threshold. This is what we see with utility satisficing utilitarianism and burden satisficing utilitarianism. We might therefore expect similar trouble at the border with ratio satisficing utilitarianism. Ratio satisficing utilitarianism, after all, also posits a relevant threshold, though the kind and location of this threshold differ from its predecessors. Fortunately, this worry is unfounded. For the existence of a threshold is not enough to pose problems: only some thresholds do. To see this more clearly, let us look back at how the thresholds of utility satisficing utilitarianism and burden satisficing utilitarianism work. In both cases, we take one dimension of what is relevant for moral permissibility—utility in the case of utility satisficing utilitarianism and burdensomeness in the case of burden satisficing utilitarianism—and set a threshold for this dimension that limits when and how alterations along this dimension can influence permissibility. This would make sense if we were interested in only a single moral permissibility-relevant dimension. However, if we take the objection from overdemandingness seriously, then this is a mistake. There are two relevant dimensions, utility and burdensomeness, both of which we need to consider. By positing a unidimensional threshold that applies to only one of these, we put on a normative eyepatch, making ourselves partly blind: we are unable to notice some alterations in a permissibility-relevant dimension, leading to a skewed picture of normative reality. This is what happens in the cases I have discussed above. If we wish to get these cases right, we need to have everything in view.

Ratio satisficing utilitarianism's threshold works quite differently. Instead of limiting the influence of one dimension, it concerns the precise relation between them. Because of this, we cannot have cases where differences in one dimension that should intuitively influence moral permissibility are inappropriately ignored by our theory. If the value of one dimension varies, it by necessity has an impact on the relation between the two dimensions and thus always has some influence. It is therefore impossible to have a case where, because of the threshold, we lose sight of one of the factors that we deem important and go astray in our moral permissibility judgments.

Certainly, there are cases where a sufficient burden-utility ratio is only barely reached or failed, and an alternative with a better outcome but accompanying higher burdensomeness is merely permissible instead of obligatory (or vice versa). This does not seem to me problematic. If we feel, on the one hand, that this alternative should be classified as obligatory, then we have a problem not with the existence of a threshold but with its location. Ratio satisficing utilitarianism is perfectly compatible with quite different views about what the relevant burden-utility ratio is (and I do not defend here any one of these views). If, on the other hand, we have general reservations about the fact that a very small difference in a normatively relevant dimension (e.g., utility) can change the deontic status of an action, then this is not specific to ratio satisficing utilitarianism. Even nonsatisficing views like classical utilitarianism imply this, after all.

7. Conclusion

Classical utilitarianism is faced with the objection from overdemandingness. One response to this objection is to offer a concessive response by modifying the utilitarian theory of moral permissibility. Typically, this concessive response comes in the form of what I call the satisficing strategy, which gives up the traditional requirement of asking agents to do the best and asks them to do only good enough. The most prominent example of this, Slote's satisficing utilitarianism, turns out to be untenable; and Chappell's burden satisficing utilitarianism, while a substantial improvement, is faced with several problems that can be traced to Chappell's commitment to an absolute burden threshold. I argue that we should give up on such an absolute threshold and turn to ratio satisficing utilitarianism, according to which moral permissibility turns on the relation between the burdensomeness of the relevant actions and the utility gained by them. Ratio satisficing utilitarianism thus turns out to be the best version of the satisficing strategy on offer.

This result should be of interest not only to utilitarians. While I have generally focused on specifically utilitarian concerns—namely, regarding the underlying axiology with its prime focus on utility—many parts of my argument apply regardless of this commitment. All theories that allow for cardinally measurable moral yields have a ratio satisficing analogue with all the structural advantages I have canvassed. Instead of comparing the generated goodness or utility, we can simply compare the relevant moral yields achieved by the respective actions, thereby tempering the demands of these other theories as well. And such a tempering might be needed, for it is not only classical utilitarianism that can ask much of agents.[37] Furthermore, if my arguments are sound, this also has implications for our understanding of the objection from overdemandingness quite generally. It is not the case that a theory can be classified as overly demanding simply because in certain situations, it asks some agent to endure a specified absolute level of burdens. Instead, we also must keep the moral stakes in clear view. Once we do this, we might, in many cases, even be able to level a reverse objection: theories that incorporate an absolute threshold of demandingness turn out to be unfavorably *underdemanding* when dispensing us of duties that, given the moral stakes, we could be expected to conform to.[38]

Friedrich-Alexander-Universität Erlangen-Nürnberg
konstantin.webkon.weber@fau.de

REFERENCES

Ashford, Elizabeth. "The Demandingness of Scanlon's Contractualism." *Ethics* 113, no. 2 (2003): 273–302.

Case, Spencer. "Rethinking Demandingness: Why Satisficing Consequentialism and Scalar Consequentialism Are Not Less Demanding than Maximizing Consequentialism." *Journal of Ethics and Social Philosophy* 10, no. 1 (2016): 1–8.

Chappell, Richard Yetter. "Willpower Satisficing." *Noûs* 53, no. 2 (2019): 251–65.

Hieronymi, Pamela. "The Force and Fairness of Blame." *Philosophical Perspectives* 18, no. 1 (2004): 115–48.

Horton, Joe. "The All or Nothing Problem:" *Journal of Philosophy* 114, no. 2 (2017): 94–104.

McElwee, Brian. "Demandingness Objections in Ethics." *Philosophical Quarterly* 67, no. 266 (2017): 84–105.

Portmore, Douglas W. *Commonsense Consequentialism: Wherein Morality Meets Rationality*. Oxford University Press, 2011.

Scheffler, Samuel. *The Rejection of Consequentialism*. Rev. ed. Clarendon Press, 1995.

37 For instance, cf. Ashford, "The Demandingness of Scanlon's Contractualism."

38 For very helpful comments on and discussion of the material, I wish to thank Gregor Betz, Inga Bones, Annette Dufner, Ufuk Özbe, Adriano Mannino, Christian Seidel, and Ugur Yilmazel. Special thanks to Dorothee Bleisch and Moritz Schulz for generous written comments on the whole manuscript.

Slater, Joe. "Satisficers *Still* Get Away with Murder!" *Ergo* 10, no. 47 (2023): 1359–77.

———. "Satisficing Consequentialism Still Doesn't Satisfy." *Utilitas* 32, no. 1 (2020): 108–17.

Slote, Michael. *Common-Sense Morality and Consequentialism*. Routledge, 1985.

———. "Satisficing Consequentialism." *Aristotelian Society Supplementary Volume* 58, no. 1 (1984): 139–63.

Sobel, David. "The Impotence of the Demandingess Objection." *Philosophers' Imprint* 7, no. 8 (2007): 1–17.

Weber, Konstantin. "Schwellenwertutilitarismus." In *Handbuch Utilitarismus*, edited by Vuko Andrić and Bernward Gesang. Metzler, 2025.

Wessels, Ulla. "Beyond the Call of Duty: The Structure of a Moral Region." In *Supererogation: Royal Institute of Philosophy Supplement*, vol. 77, edited by Christopher Cowley. Cambridge University Press, 2015.

York, Kyle. "Moral Demandingness and Modal Demandingness." *Journal of Ethics and Social Philosophy* 28, no. 2 (2024): 151–70.

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY
http://www.jesp.org
ISSN 1559-3061

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

Funding for the journal has been made possible through the generous commitment of the Division of Arts and Humanities at New York University Abu Dhabi.

جامعة نيويورك أبوظبي
NYU ABU DHABI