

JOURNAL *of* ETHICS  
& SOCIAL PHILOSOPHY

VOLUME XV · NUMBER 1

*April 2019*

ARTICLES

- 1 Parental Partiality and Future Children  
*Thomas Douglas*
- 19 Prioritarianism: A (Pluralist) Defense  
*Shai Agmon and Matt Hitchens*
- 43 Reasons: Wrong, Right, Normative, Fundamental  
*Errol Lord and Kurt Sylvan*

DISCUSSIONS

- 75 Helping Buchanan on Helping the Rebels  
*Daniel Weltman*
- 85 Is There Value in Keeping a Promise? A Response  
to Joseph Raz  
*Crescente Molina*

The *Journal of Ethics and Social Philosophy* (ISSN 1559-3061) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge and is available to all readers under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license. Funding for the journal has been made possible through the generous commitment of the Gould School of Law and the Dornsife College of Letters, Arts, and Sciences at the University of Southern California.

The *Journal of Ethics and Social Philosophy* aspires to be the leading venue for the best new work in the fields that it covers, and it is governed by a correspondingly high editorial standard. The journal welcomes submissions of articles in any of these and related fields of research. The journal is interested in work in the history of ethics that bears directly on topics of contemporary interest, but does not consider articles of purely historical interest. It is the view of the associate editors that the journal's high standard does not preclude publishing work that is critical in nature, provided that it is constructive, well-argued, current, and of sufficiently general interest.

*Editor*

Mark Schroeder

*Associate Editors*

Dale Dorsey

James Dreier

Julia Driver

Anca Gheaus

Tristram McPherson

Colleen Murphy

Hille Paakkunainen

*Discussion Notes Editor*

Kimberley Brownlee

*Editorial Board*

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

*Managing Editor*

David Clark

*Copyeditor*

Susan Wampler

*Typesetting*

Matthew Silverstein



## PARENTAL PARTIALITY AND FUTURE CHILDREN

*Thomas Douglas*

PARENTS ARE TYPICALLY partial to their own children. They typically treat their own children more favorably, on certain dimensions, than other similarly placed individuals. For example, most parents invest more emotional and financial resources in sheltering, nourishing, educating, and entertaining their children than they invest in others who are equally in need of these goods, and would benefit equally from them. Moreover, most parents take themselves to be justified in engaging in such partiality, and most philosophical theorists of partiality agree.

Less commonly noted is that parents can also be partial toward their *future* children—children who will, but do not yet, exist. They can, for example, set aside resources for their own future children in preference to investing them in others who could make similar use of those goods. Consider:

*A&B* plan to, and will in fact, have a child together, but they have not yet conceived. They win £10,000 in the lottery and are considering whether to set it aside for their future child's education or donate it to a charity that helps to provide education to impoverished children in Africa. They decide to set it aside for their future child; they simply care more about *their* child's education.

Like parental partiality, such “pre-parental partiality” plausibly stands in need of justification.<sup>1</sup> For one thing, it is unclear whether and how pre-parental partiality—like partiality more generally—can be reconciled with the plausible

1 It might be denied that there is any partiality in this case if the child *A&B* will have (i) would not have existed had they not engaged in the partiality, and (ii) is not benefitted by being brought into existence. Perhaps partiality necessarily involves *benefitting* the object of the partiality. However, regardless of whether the case of *A&B* is aptly described as a case of partiality, it bears sufficient structural similarities to paradigmatic cases of partiality that it seems worthwhile to consider how justifications for partiality might bear upon it, which is what I will do here. I will continue to refer to cases like that of *A&B* as cases of partiality, though nothing will turn on whether they are aptly described.

thought that everyone matters equally.<sup>2</sup> For another, pre-parental partiality can have troubling distributive effects.<sup>3</sup> For instance, if economically well-off parents direct their economic resources preferentially to their own future children, those children are likely to wind up better off, economically, than the (present or future) children of parents who are less well off, or who allocate their economic resources less partially.

It thus seems legitimate to demand a moral justification for pre-parental partiality. However, unlike parental partiality, pre-parental partiality has not enjoyed the sort of philosophical attention that might be expected either to furnish or to foreclose such justification. Though the *possibility* of pre-parental partiality has been noted in ethical discussions of reproductive technologies, its *justifiability* has attracted little comment.<sup>4</sup>

So what should we make of pre-parental partiality? What moral assessment does it merit? Were we to take commonsense morality as our guide, we would, I think, be obliged to take a highly permissive stance: commonsense morality grants parents substantial leeway to be partial to their current children, and it is, I think, just as permissive of pre-parental partiality. However, I will argue that *philosophical justifications* of parental partiality do not carry over so easily to the pre-parenthood case. Insofar as the reasons in favor of the two kinds of partiality are established through appeal to existing philosophical justifications, the reasons in favor of pre-parental partiality will typically be weaker than reasons in favor of parental partiality. Thus, either reasons in favor of pre-parental partiality *are indeed* typically weaker than reasons in favor of parental partiality—suggesting, contrary to common sense, that pre-parental partiality will be permissible in a narrower range of circumstances—or the literature on partiality is importantly incomplete: it does not capture all of the reasons in favor of partiality in play in these cases.<sup>5</sup>

2 See Pettit and Goodin, “The Possibility of Special Duties”; Keller, *Partiality*, ch. 1. Keller calls this the “puzzle of partiality.”

3 For discussion of distributive concerns regarding parental and other varieties of partiality toward existing people, see Scheffler, “Relationships and Responsibilities”; Brighouse and Swift, “Legitimate Parental Partiality.”

4 For discussion of pre-parental partiality in relation to reproductive technologies, see, for example, Lillehammer, “Reproduction, Partiality and the Non-Identity Problem”; Elster, “Procreative Beneficence”; Douglas and Devolder, “Procreative Altruism: Beyond Individualism in Reproductive Selection”; Petersen, “On the Partiality of Procreative Beneficence”; Lewandowski, “Parents, Special Obligations and Reproductive Genetics.” For discussion of the justifiability of pre-parental partiality, see Lillehammer, “Reproduction, Partiality and the Non-Identity Problem”; Lewandowski, “Parents, Special Obligations and Reproductive Genetics.”

5 I assume here that the reasons *against* partiality are similar in parental and pre-parental cas-

Throughout, I will limit myself to cases in which, not only does the child in question not yet exist, she has also not yet been conceived.<sup>6</sup> These are, I think, the cases in which the task of justifying partiality is most challenging. I henceforth use the terms “pre-parent” and “prospective parent” interchangeably to refer to *pre-conception* prospective parents.

### 1. THREE TYPES OF JUSTIFICATION

In identifying candidate justifications for pre-parental partiality, I will draw on the literature on the justification of *parental* partiality, and that on the justification of partiality more generally. In those literatures, three broad types of justification can be found.

Justifications of the first type—*telic* justifications—appeal to some value that is (expectably) realized by the partiality.<sup>7</sup> In the case of partiality between family members, friends, or romantic partners, the value invoked is often love.<sup>8</sup> John Cottingham’s justification of “philophilic partialism”—partiality within loving relationships—provides a clear example:

The justification of philophilic partialism is ... extremely simple. If I give

- 
- es. In both, the chief concern is that partiality could have unjust or otherwise undesirable distributive effects, and I see no reason to suppose that this problem would be greater in the one case than the other.
- 6 Some will hold that the point at which the child comes into existence is the point of conception, and I do not mean to contradict that view here.
  - 7 I henceforth omit the “(expectably).” Influential examples of telic justifications have been offered by John Cottingham and by Harry Brighouse and Adam Swift. See Cottingham, “Partiality, Favouritism and Morality”; Brighouse and Swift, “Legitimate Parental Partiality” and *Family Values*. Scheffler (“Relationships and Responsibilities”) defends a justification according to which partiality contributes (in his account, constitutively) to a relationship that the agent of the partiality *has net, noninstrumental reasons to value*. This justification is equivalent to a telic justification if we have net, noninstrumental reasons to value something just in case the thing in fact *has value*. However, even if Scheffler’s justification is not equivalent to a telic justification it has, I believe, the same implications for the justification of pre-parental partiality. Thus, I will not consider it separately.
  - 8 Such love-based justifications for parental partiality should not be confused with justifications *for* love. The latter, if successful, justify parental partiality, insofar as such partiality is a constituent of love. Moreover, they may derive some of their justificatory force *from* an appeal to love. For example, Niko Kolodny holds that part of the justification for loving someone is that one has had a loving relationship with that person in the past; see Kolodny, “Love as Valuing a Relationship.” However, justifications for love—and thus for any partiality that is a constituent of love—need not attribute any normative role to (other aspects of) love in justifying the partiality. They may, for example, invoke some third factor that directly justifies both the partiality and the (other aspects of) love.

no extra weight to the fact that this is my love, my friend, my spouse, my child, if I assess these people's needs purely on their merits (in such a way as an impartial observer might do), then that special concern which constitutes the essence of love and friendship will be eliminated. Partiality to loved ones is justified because it is an essential ingredient of one of the highest human goods.<sup>9</sup>

The thought here is that partiality is justified by virtue of its contribution to realizing valuable love, or some particular kind of valuable love. On Cottingham's picture, the contribution of partiality to love is constitutive; treating a person partially is part of what it is to love that person. But we can equally imagine a view on which partiality is justified by its *causal* contribution to the development of love. Perhaps partiality tends to elicit feelings of affection for the agent of the partiality, and perhaps these feelings partly constitute love.

Justifications for partiality of a second type—*deontic* justifications—invoke some relationship between the agent and recipient of the partiality, and hold this to provide non-value-based (henceforth, “deontic”) reasons of partiality. It is plausible, for instance, that a contractual relationship could play this role; that you and I have agreed to share my lottery winnings may give me a reason to treat you partially in disbursing those winnings, and that reason is plausibly (though not uncontroversially) independent of any value that will be realized by my upholding the agreement.<sup>10</sup> Contractual relationships are, however, not the only relationships that have been thought to play such a role. For example, some have held that being in a (certain kind of) loving relationship with another may give one a deontic reason to treat that other partially.<sup>11</sup> Similarly, some have held that being responsible for creating a person may give one such a reason, regardless of whether it implies the existence of an implicit contract.<sup>12</sup>

Finally, what I will call *hybrid* justifications of partiality hold that reasons of partiality are ultimately telic reasons—reasons to realize value—but these reasons are enabled or strengthened by the relationship between the agent and re-

9 Cottingham, “Partiality, Favouritism and Morality,” 369.

10 I am here entertaining the possibility that some reasons are not reasons to realize value, however I do not mean to commit myself to this view. If there are no such reasons, so much the worse for this class of justifications for partiality.

11 See, for an example of this approach, Jeske, “Families, Friends, and Special Obligations.”

12 See, for this approach and variants thereof, Rakowski, *Equal Justice*, 153–55; Steiner and Valentyne, “Libertarian Theories of Intergenerational Justice”; Olsaretti, “Liberal Equality and the Moral Status of Parent-Child Relationships”; Ferracioli, “Procreative-Parenting, Love's Reasons and the Demands of Morality.” Anca Gheaus argues that it is *gestating* a person that generates reasons of partiality; see Gheaus, “The Right to Parent One's Biological Baby.”

recipient of the partial treatment. Simon Keller defends a view of this kind.<sup>13</sup> He argues that reasons of partiality are reasons to promote the good of the recipient of the partiality. Thus, for example, your child's value is the source of your reasons to treat your child partially. But of course, other children are also valuable—perhaps equally so—yet you do not have similarly strong reasons in relation to them. This, on Keller's view, is because features of your relationship with your child enable your child's value to generate special reasons for you, while no similar enabling condition is present in relation to other children.

Philip Pettit and Robert Goodin's justification for partiality within special relationships can also be understood as a hybrid justification.<sup>14</sup> On their view, duties of partiality are a function of both value and responsibility; they are duties to realize states of affairs that are valuable *and for which one is responsible*. Though Pettit and Goodin do not put it in these terms, we can think of one's responsibility for a state of affairs as *enabling* the value of that state of affairs to generate (especially strong) reasons.

It might be thought that there is a fourth class of justifications that needs to be considered: "two-level" justifications, according to which people have reasons to treat some others partially because they would be assigned such reasons under the optimal scheme for distributing our more general, agent-neutral moral obligations across people.<sup>15</sup> However, even if reasons of partiality have this deep source, they will have a surface structure that falls into one of the categories that I mentioned above. Thus, we do not need to consider such two-level justifications for partiality as a separate category. We can simply acknowledge that justifications of the three varieties I have mentioned—telic, deontic, and hybrid—may themselves be subsumed within a two-level justification of this sort. Similar thoughts apply to sophisticated forms of consequentialist justification, according to which one has most reason to do whatever act has the agent-neutrally best consequences, but one may also have most reason to adopt a decision-making procedure that involves deciding *as if* one had different reasons, perhaps including reasons of partiality.<sup>16</sup> We can regard telic, deontic, and hybrid justifications as three different models of the reasons that one would, under the decision-making procedure one ought to adopt, take oneself to have.

13 See, Keller, *Partiality*.

14 Pettit and Goodin, "The Possibility of Special Duties." See also Goodin, "What Is So Special about Our Fellow Countrymen?"

15 For an example of this approach, see Goodin, "What Is So Special about Our Fellow Countrymen?"

16 See, for the classic statement of this sort of view, Railton, "Alienation, Consequentialism, and the Demands of Morality."

The important point is that, regardless of whether these three different types of justification for partiality are subsumed under a deeper agent-neutral justification, they have different implications for the justification of *pre-parental* partiality. In the next two sections, I consider whether deontic justifications might be able to justify pre-parental partiality, either directly (section 2) or indirectly (section 3). In section 4, I turn to consider telic justifications. I do not explicitly examine hybrid justifications for pre-parental partiality, but I hope it will be clear how my arguments would apply to such justifications. I believe that they face all of the problems that I raise for telic justifications, as well as some of those that I raise for deontic justifications.

## 2. DEONTIC JUSTIFICATIONS

It may seem that, if being in a certain kind of relationship with someone is to give one a deontic reason to be partial, that relationship must already exist. Consider: if being in a contractual relationship with someone is to give one a reason to fulfill the terms of the contract—and thus, perhaps, a reason to be partial—one must already have formed the contract. The mere fact that you and I *will* agree to send one another comments on our papers does not yet give me any reason to send you such comments. At least, it does not give me the ordinary kind of contractual reason to do so. (I return in the next section to consider the possibility that it might give me some other kind of reason.) This raises the question: do any of the relationships that have been thought to provide reasons of partiality already exist between a pre-parent and her future child?<sup>17</sup>

Above, I mentioned three kinds of relationship that have been thought to generate deontic reasons of partiality: contractual relationships, procreative relationships, and loving relationships. It seems clear that pre-parents do not already stand in a *contractual* relationship with their future child. Perhaps parents have normally agreed—explicitly or implicitly—to provide certain kinds of support to their existing children. However, the most plausible points at which that contract is formed are the point of adoption (in the case of adopted children), the point at which the choice is made *not* to adopt a child out, or the point of voluntary conception. Whichever of these points one chooses, *pre-parents*—that is, preconception prospective parents—have not yet formed the contract.

It is even clearer that preconception prospective parents do not already stand in a procreative relationship with their future child; the process of creation has not yet begun. But could a prospective parent already stand in a *loving* relation-

17 Niko Kolodny raises a restricted version of this question in “Which Relationships Justify Partiality?” 65–66.

ship with her future child? This is perhaps more plausible. After all, diachronic love with people who do not currently exist does seem possible. It seems possible for a person to love her long-lost grandparent, for instance.

Nevertheless, we can doubt that a pre-parent can enjoy a loving relationship with her future child. For a loving relationship to exist between two individuals, it is plausibly necessary that at least one party to the relationship love the other.<sup>18</sup> But clearly a child cannot love his parent before the child exists. And we might also doubt whether a parent can love a child before the child exists. This is because loving someone plausibly requires a kind of attachment to a particular individual. We can see this by considering relationships that seem otherwise to mimic love but lack this attachment. Imagine a parent who pours affection on her child, is highly partial toward her child, derives great pleasure from the flourishing of that child, but is completely indifferent to whether that child is replaced by another. We would not want to say that this parent loves her child, and the obvious way of accounting for this would be to invoke the view that love requires attachment to a particular individual.

It is, however, difficult to see how a prospective parent could have any attachment to her future child as a particular individual. First, in typical cases, the identity of the future child could be influenced by the partial action whose justification is in question or by other choices subsequently made by the parents or others. There is thus a sense in which, prior to that action, the identity of the future child has not yet been fixed: it remains within the sphere of human choice. Second, even in cases where the identity of the future child has been fixed, in the sense that it will not be altered by the parents' or anyone else's choices, the prospective parents will not have any information about their particular future child that would allow them to fix on that child as a particular individual.<sup>19</sup> Indeed, they will typically know nothing that uniquely picks out the particular future

18 I use "love" here and throughout to refer only to what is sometimes called "interpersonal love," as distinct from the kind of love that a person can have for an inanimate object.

19 On one influential account of love, David Velleman's, love consists not in the presence of any positive attachment, but rather in the letting down of certain emotional defenses. Love "arrests our tendencies toward emotional self-protection from another person" ("Love as a Moral Emotion," 361). Note, however, that Velleman does still think of love as directed toward a particular individual who is viewed as incomparable to and irreplaceable by others (see especially 364–68). Thus love, on Velleman's account, still requires that the lover can latch on to the object of his love as a particular person. Indeed, Velleman makes this explicit, noting that "unless we actually see a person in the human being confronting us, we won't be moved to love; and we can see the person only by seeing him in or through his empirical persona" (371).

child that they will have, distinguishing that child from other possible children that they might have had.<sup>20</sup>

In any case, even if prospective parents *can* form such an attachment, they typically do not. This can be seen by considering cases in which a third party acts so as to alter the identity of the future child. Suppose that Arama and Bram are due to undergo a fertility treatment in which one of Arama's eggs will be fertilized by one of Bram's sperm *in vitro*. You phone Bram just as he is about to leave for the fertility clinic to give the sperm sample. As a result, he gives the sample fifteen minutes later than would otherwise have been the case, with the result that Arama's egg is fertilized by a different sperm. You have altered the identity of Arama and Bram's future child, yet it seems doubtful that they would feel that they have suffered any kind of loss. They would regard their new, post-phone-call future child as a perfect replacement for the pre-phone-call child. This indicates that they had no attachment to the particular future child that they would otherwise have had.

I have been suggesting that a loving relationship normally does not exist between a pre-parent and her future child because neither the pre-parent nor the future child loves the other. However, even if a loving relationship does exist between a pre-parent and her future child, there are reasons to doubt whether this love could be as normatively significant as parent-child love—reasons to doubt, that is, that it could generate similarly strong deontic reasons of partiality. It is plausible that at least some of the normative significance of parent-child love derives from a feature of the parent-child relationship that a relationship between a pre-parent and her future child cannot have: intimacy.

Parent-child relationships typically involve intimacy—by which I mean repeated positive, personal interactions that make each vulnerable to the other. Moreover, the intimacy of paradigmatic parent-child relationships, and the closest romantic relationships and friendships, is plausibly part of what gives these loving relationships their peculiarly great normative force.

To see this, consider a case in which a previously intimate relationship loses its intimacy. Suppose Cedric and Danele were close friends who saw each oth-

20 I am not here claiming, *contra* Harry Frankfurt, that one must appreciate the value of another, or have significant information about their personal characteristics, in order to love that other. I am claiming only that one has to have enough information to fix on the other as a particular person. See Frankfurt, *The Reasons of Love*, esp. 39. I am also not suggesting that there are *never* cases in which pre-parents would have enough information to fix on a particular child. We could think here of a case in which the pre-parents have already selected the egg and sperm that will be used to create their future child, and already committed to using a genetic technology to “design in” some unique genetic characteristics to that child. I thank an anonymous reviewer for pressing me to consider cases of this sort.

er and engaged in highly personal and revealing conversations multiple times per day, but who then became separated from one another, and completely lost contact for a number of years, though their deepest positive attitudes toward one another remained unchanged. Several years after this separation, Cedric and Danele happen upon each other on the street, and Cedric asks Danele to do him a significant favor—to spend the weekend helping him move apartments, for example. It is plausible that some kind of loving relationship still exists between the two and that this gives Danele some reason to help Cedric with his move. However, it is also plausible that this reason is significantly weaker than it would have been had Cedric asked for the help at a time when the two were still in regular and close contact. This suggests that the loss of what I am calling intimacy has diminished the normative significance of the friendship. It is, moreover, possible to give a plausible explanation for why the loss of intimacy causes this loss of significance: intimacy arguably produces a web of mutual expectations and implicit commitments that gives the parties to the intimate relationship special reasons to help and care for one another.

These thoughts indicate that, even if it is possible for a loving relationship to exist between a pre-parent and her future child, this relationship will be in one respect less normatively significant than paradigmatic parent-child relationships, since a pre-parent clearly cannot enjoy intimacy with her future child. Since it is difficult to see any countervailing respect in which pre-parent/future-child love, if it can exist, would be *more* normatively significant than parent-child love, this suggests that it will, overall, generate less strong deontic reasons.

I have been arguing that contractual, procreative, and loving relationships either fail to establish deontic reasons of pre-parental partiality, or typically establish only weaker deontic reasons in support of such partiality than in support of parental partiality. Are there other relationships that might generate deontic reasons of pre-parental partiality that are just as strong as our reasons for parental partiality? Perhaps there are. After all, a pre-parent and her future child do stand in *some* significant relationships to one another. In cases of genetic parenthood, they stand in a rather close genetic relationship. They also stand in the relationship of extended family co-membership. It might be thought that one or other of these relationships generates deontic reasons to be partial. Moreover, since these same relationships exist between a current parent and her existing child, it might seem that these relationships will generate reasons of partiality that are equally strong in the parental and pre-parental cases.

It is, however, not plausible that an appeal to these relationships can fully account for the reasons in favor of ordinary parental partiality. To see this, note that genetic relationships and co-memberships of social groups are sometimes

invoked to justify partiality between members of the same racial, species, or national groups. And though many would argue that some partiality between the members of national and species groups is justified by the existence of these relationships, it is not plausible that the reasons to engage in such forms of partiality are as strong as those to engage in parental partiality. Similarly, some might hold that genetic parents who are not social parents have reasons of partiality in respect of their genetic children, and that family co-members whose relationship is more distant than a parent-child relationship have reasons of partiality in respect of each other, but it is not at all plausible that these reasons are as strong as the reasons of partiality that social parents have with respect to their social children. This suggests that there are further reasons to engage in parental partiality—reasons above and beyond those provided by genetic relationships and social group (including family) co-membership. Plausibly, some of these reasons are based on the existence of a contractual, procreative, or loving relationship in parental cases. But, if my arguments above are sound, these reasons do not obtain in pre-parental cases, or are less forceful in such cases than in cases of parental partiality. Thus, even if the reasons of partiality generated by genetic relationships and social group co-membership are equally strong in pre-parental and parental cases, we remain without any basis for thinking that reasons of pre-parental partiality are, *all things considered*, as strong as our reasons of parental partiality.

### 3. ENSURING MORAL COMPLIANCE

I suggested above that contractual reasons of the ordinary kind only obtain once the contractual relationship is in place, and likewise for deontic reasons deriving from procreative or loving relationships. But perhaps these relationships could generate other kinds of deontic reasons that obtain even before the relevant relationship exists. Let us again take the case of contractual reasons as illustrative. Though ordinary contractual reasons only obtain once the contract has been formed, those reasons plausibly give rise to derivative reasons that obtain in advance of the contract. Suppose I know that I will agree to give you comments on your draft paper. An ordinary contractual reason to give you these comments is not yet in place, but it is plausible that some derivative reasons *are* already in place. I may already have reasons not to do things that make me less likely to live up to my agreement. For example, I may have reasons not to promise others comments on their papers, knowing that I will not be able to comment on more than one paper.<sup>21</sup> More generally, it might seem that we each have moral reasons

<sup>21</sup> The existence of these derivative reasons need not depend on my already having the

not to prevent our compliance with our future contractual reasons, and, indeed, with our future moral reasons more generally. This raises the question of whether pre-parental partiality might be justified on the ground that failing to engage in such partiality could be expected to prevent one's compliance with one's future reasons of parental partiality, whatever the source of those future reasons.

Let us begin to explore this possible justification for pre-parental partiality by seeing how it plays out in the case of *A&B*, with which I began. In the original version of the case, *A&B* set aside their lottery winnings for the education of their future child rather than donating them to an African educational charity. Suppose they had instead taken the opposite course, donating their winnings to the African charity. Would they then have prevented their future compliance with reasons of partiality?<sup>22</sup>

One reason to doubt that they would becomes apparent when we attend to the nature of the reasons of partiality that *A&B* will have once their child exists. We can distinguish two importantly different possibilities. One possibility is that *A&B*'s reasons of partiality, once their child exists, will be reasons to (expectably) confer some benefit (or prevent some harm) that is independent of *A&B*'s motives for conferring (preventing) it.<sup>23</sup> For example, perhaps *A&B* will have reason to confer some given level of education on their future child. Call reasons of this sort *reasons to benefit*.

A second possibility is that *A&B*'s reasons of partiality are reasons to give certain *consideration* or *significance* to their child in their decision making about whom to benefit and to what degree. For example, perhaps *A&B* will have reason to give their child greater consideration than other children in deciding how to allocate their resources. Call reasons of this sort *reasons to consider*. Unlike reasons to benefit, which are reasons to bring about particular states of affairs, reasons to consider are reasons to deploy particular procedures in deciding which states of affairs to bring about.<sup>24</sup> I believe that at least some of the reasons of

---

*intention* to form the contract. It may be enough that the contract will be formed, or that I reasonably believe it will be formed.

22 Note that, if "reason to" implies "can," then, in preventing future parental partiality, *A&B* may have prevented themselves from falling under reasons of partiality, rather than preventing their compliance with those reasons. In this case, we will need to understand their putative pre-parental reasons not as reasons to ensure their compliance with reasons that they *will* have to engage in parental partiality, but as reasons to ensure their compliance with reasons that they *would have had* had they acted otherwise.

23 This benefit could be absolute, or relative to the benefit enjoyed by others.

24 It is plausible that some reasons are neither simply reasons to benefit nor simply reasons to consider, but are rather mixed reasons: for example, reasons to bring about certain benefits for certain reasons, or reasons to bring about benefits that constitute benefits in part in

partiality that parents have with respect to their children are reasons to consider rather than reasons to benefit.

Suppose that Edith is a moderately well-off parent who believes that parental partiality is unjustified and strives to allocate no more time, effort, or money to her own children than to other equally needy children with whom she has contact. However, as it happens, she very rarely comes into contact with other children, and when she does, they are typically less needy than her own child. Thus, she ends up allocating far more time, effort, and money to her own child than to others; in fact, the share of her resources that go to her own child relative to others is typical for other similarly well-off parents. Edith's neighbor Femke is similarly well-off, and is also a parent. Like Edith, Femke allocates to her child a share of her time, effort, and money that is typical of similarly well-off parents. However, she does this for the usual reason: because she gives priority to the interests of her own child in her practical deliberation.

It seems clear that there is something defective about Edith's parenting that is not defective about Femke's parenting. Yet both parents apparently benefit their children to the same degree, both in absolute terms and relative to other children. The most straightforward way to explain the defect in Edith's parenting is, I believe, to suppose that, in addition to reasons to benefit their children, parents have reasons to give certain consideration to their children in practical decision making. The problem with Edith's parenting is that she does not give sufficient consideration to her own child and thus fails to fully comply with these reasons.<sup>25</sup>

This has important implications for the justification of pre-parental partiality in the case of A&B. We were supposing that A&B's pre-parental partiality in setting aside their lottery winnings might be justified by its tendency to promote compliance with their future reasons of parental partiality. But to the extent that A&B's future reasons of parental partiality will be reasons to consider rather than reasons to benefit, it is not clear how this justification could succeed, for A&B's setting aside the £10,000 is plausibly irrelevant to the degree of consideration that they will subsequently give to their child in their decision making. Had A&B donated their lottery winnings to the African charity rather than setting them aside, they would have restricted the maximum possible educational benefit that they could subsequently confer on their future child. But it is not at all clear that they would have diminished the degree of consideration that they will

---

virtue of the reasons for which they are conferred. However, for my purposes here, we can ignore reasons of these kinds.

25 As a conceptual matter, it also seems to me doubtful that we could aptly describe Edith as treating her child partially; partiality necessarily involves giving special consideration to the target of the partiality.

give to their child in allocating whatever resources they will in fact have. They will be just as free, and plausibly at least as likely, to give special consideration to their future child in distributing their wealth if they give the £10,000 away; they will simply have less wealth to distribute.

All of this suggests that A&B's derivative reasons to engage in pre-parental partiality will be less powerful than their later reasons to engage in parental partiality. Once A&B have their child, they will plausibly have both reasons to benefit and reasons to consider in respect of that child, but only the reasons to benefit give rise to derivative reasons to engage in pre-parental partiality.

There are, no doubt, *some* cases in which both reasons to benefit and reasons to consider generate derivative reasons to engage in pre-parental partiality, since pre-parental partiality contributes to future compliance with both sorts of reason. Suppose that Gael expects to have, and love, a child in the future. He is considering whether to cultivate a disposition to be highly partial to his future child, or a disposition to be an unusually impartial parent—one who gives little more weight to his own child than to others. If Gael chooses to cultivate the impartial disposition, he will diminish his future compliance with reasons of partiality regardless of whether those reasons are reasons to benefit or reasons to consider. This is a rather special case, however. In typical cases where pre-parental partiality is an option, failing to engage in that partiality will not frustrate compliance with future reasons to consider. Thus, reasons of pre-parental partiality will, on the present approach, typically be weaker, all things considered, than the reasons of parental partiality from which they derive.<sup>26</sup>

#### 4. TELIC JUSTIFICATIONS

Let us turn now to consider telic justifications. As candidate justifications for pre-parental partiality, telic justifications have an advantage over deontic justifications: the values they invoke can directly generate reasons even if the values do not yet obtain. On telic justifications, it is enough that the partial treatment will (expectably) realize the value at some point in the future. A telic justification for pre-parental partiality could succeed even if the value(s) invoked will obtain only once the child has come into existence.

<sup>26</sup> Of course, reasons of pre-parental partiality could still be strong even if they derive only from reasons to benefit. After all, some may hold that reasons to benefit are powerful in parental cases—perhaps sufficiently strong to, on their own, establish *obligations* to engage in parental partiality. Nevertheless, the result that reasons of pre-parental partiality are typically weaker than reasons of parental partiality is significant, since it suggests that pre-parental partiality will be permissible, and obligatory, in a narrower range of cases than parental partiality.

Telic justifications can be broadly split into two classes. Justifications of the first sort invoke some particular value and maintain that this value can best, or only, be realized through partiality. Cottingham's abovementioned justification, according to which partiality constitutively contributes to love, exemplifies this approach. By contrast, justifications of a second kind invoke no special value. Rather, they try to show that, in treating certain individuals near and dear to us favorably, we will generally produce more of *whatever* is good for people. This is because we are better placed to advance the interests of those near and dear to us than to advance the interests of others.<sup>27</sup> At least part of the story here is that we are normally in a particularly good epistemic position with respect to those near and dear to us: we are particularly good at identifying their interests.

Telic justifications of the second sort do not seem promising as a basis for establishing reasons of pre-parental partiality as strong as those of parental partiality. This is because at least part of the case for partiality on this view—the part that invokes epistemic considerations—does not carry over in full from the parental to the pre-parental case. Though parents typically are in a strong position to know what will advance the interests of their existing children, they are normally in a much weaker epistemic position with respect to their future children. Though pre-parents will normally know some general facts about the kind of social environment their future children are likely to occupy, and may be able to make some reasonable assumptions about the kinds of genetic predispositions they are likely to possess, they will not have the sort of detailed knowledge regarding a child's individual needs and proclivities that is typical of parents.

Let us focus, then, on telic justifications of the first sort—justifications invoking particular values that can best or only be realized through partiality. As we have seen, one value sometimes invoked in such justifications is love. Another is the value of completing (or advancing) parental projects.<sup>28</sup> It might seem that there is generally value in advancing one's projects, at least when those projects are not morally prohibited, and many people have projects that can be most effectively pursued through parental partiality.

Each of these values may, it seems to me, justify some pre-parental partiality. Consider forms of pre-parental partiality that help secure a stable and rich social environment for the future child. It is quite plausible that such partiality helps foster forms of emotional intelligence in the child that are conducive to the development of parent-child love. It is also plausible that some forms of pre-pa-

27 See, for an example of this approach, Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection."

28 See, for example, Williams, "Persons, Character and Morality"; Wolf, "Morality and Partiality"; Stroud, "Permissible Partiality, Projects, and Plural Agency."

rental partiality contribute significantly to the advancement of the pre-parents' projects. Suppose, for instance, that a couple have adopted the project of raising a child who has every chance of becoming a music star, or at least of fulfilling her potential for such. In order to most effectively pursue this project, these pre-parents may need to set aside significant funds for private music tuition *before* the child is conceived.

Still, there are reasons to think that love- and project-based telic justifications will provide weaker support for pre-parental partiality than for parental partiality. Consider first *love-based* telic justifications. It is plausible that one way in which *parental* partiality typically contributes to love is by expressing parental attachment to the child. The expression of such attachment may be partly constitutive of love, but even if it is not, it plausibly contributes causally to love. I have already argued, however, that pre-parents typically do not have any attachment to their particular future child, so it is not clear how pre-parental partiality could express such an attachment. This suggests that, in at least this one respect, pre-parental partiality will be less conducive to the realization of love than parental partiality.

Consider next *project-based* telic justifications. It is plausible that our reasons to advance our personal projects are generally weaker in the early stages of the project than they are in the later stages. Thus, for example, we would normally think that a person has weaker project-based reasons to advance her project of writing a novel when she is only one page in than when she only has one chapter left to write. In the early stages, we would think there is little to be lost in sacrificing the project in order to, say, adopt another one; in the later stages, we might think that any such substitution of projects would involve a significant loss. This point is relevant to the case of pre-parental partiality since pre-parents are typically only in the early stages of their parental project. This suggests that their project-based reasons will typically be weaker than those of parents.

I contend, then, that both love- and project-based telic reasons of partiality will typically be weaker in pre-parental cases than in parental cases. However, even if this is not so, it is doubtful that an appeal to such reasons could show that, all things considered, the reasons that support parental partiality are normally as strong as those supporting parental partiality. This is because, as with the appeals to genetic relationships and social group co-membership discussed in section 2, it is doubtful that either of these justifications can fully account for *parents'* reasons to treat their children partially: it is doubtful that the *only* reasons of parental partiality are project-based and love-based telic reasons. This can be seen by considering a case in which, due to estrangement, a parent has no prospect of realizing a loving relationship with her child, and the child plays no

role in the parent's projects.<sup>29</sup> It seems hardly plausible that such a parent would have *no* reasons of partiality in respect of her child in such a case. We would, for example, surely still think that, were the estranged child to fall into desperate need, the parent would have *some* special reason to help him. There seem, then, to be reasons of parental partiality besides telic reasons to realize love or advance parental projects. If, say, those further reasons include the reasons invoked by deontic contract- or procreation-based justifications, and my arguments above are sound, then these other reasons do not fully carry over to pre-parental cases. In that case, the conclusion that reasons of pre-parental partiality are *overall* weaker than those of parental partiality will stand.

### 5. CONCLUSIONS

I have now considered three kinds of justification that might be offered for pre-parental partiality.

Justifications of the first kind appeal directly to deontic reasons that arise from certain kinds of relationships. I argued that the relationships normally invoked to establish these reasons—contractual, procreative, and loving relationships—do not obtain between prospective parents and their future children, or obtain only in a less normatively significant form. By contrast, the genetic and familial relationships that do exist between prospective parents and their future children generate only weak deontic reasons of partiality, and do not fully account for reasons of parental partiality.

Justifications of the second kind hold that prospective parents *will* have reasons to engage in parental partiality once their child exists. They seek to justify pre-parental partiality as a means to ensuring compliance with those reasons. I argued that such derivative reasons of pre-parental partiality will typically be weaker than the reasons of parental partiality from which they derive: in most cases, pre-parental partiality will not increase compliance with some reasons of parental partiality—those that are reasons to *consider*.

Finally, on justifications of the third kind—telic justifications—partiality is justified by the values that it (expectably) realizes. I argued that the values most commonly invoked in support of parental partiality—the values of project advancement and love—will typically provide weaker support for pre-parental partiality than for parental partiality. I also questioned whether these telic reasons exhaust our reasons in favor of parental partiality.

The upshot of my arguments is this: existing justifications for partiality

29 For further reasons to doubt that project-based reasons exhaust reasons of parental partiality, see Keller, *Partiality*, ch. 2.

typically establish weaker reasons in support of pre-parental partiality than in support of parental partiality. Thus, either these existing justifications do not fully account for our reasons of parental partiality, or our reasons to engage in pre-parental partiality are indeed typically weaker than our reasons to engage in parental partiality. If the latter, we can expect that, contrary to common sense, pre-parental partiality will be morally permitted, and indeed morally required, in a narrower range of circumstances than parental partiality.<sup>30</sup>

University of Oxford

thomas.douglas@philosophy.ox.ac.uk

#### REFERENCES

- Brighouse, Harry, and Adam Swift. *Family Values: The Ethics of Parent-Child Relationships*. Princeton: Princeton University Press, 2014.
- . “Legitimate Parental Partiality.” *Philosophy and Public Affairs* 37, no. 1 (Winter 2009): 43–80.
- Cottingham, John. “Partiality, Favouritism and Morality.” *Philosophical Quarterly* 36, no. 144 (July 1986): 357–73.
- Douglas, Thomas, and Katrien Devolder. “Procreative Altruism: Beyond Individualism in Reproductive Selection.” *Journal of Medicine and Philosophy* 38, no. 4 (August 2013): 400–19.
- Elster, Jakob. “Procreative Beneficence – Cui Bono?” *Bioethics* 25, no. 9 (November 2011): 482–88.
- Ferracioli, Laura. “Procreative-Parenting, Love’s Reasons and the Demands of Morality.” *Philosophical Quarterly* (forthcoming).
- Frankfurt, Harry. *The Reasons of Love*. Princeton: Princeton University Press, 2004.
- Gheaus, Anca. “The Right to Parent One’s Biological Baby.” *Journal of Political Philosophy* 20, no. 4 (December 2012): 432–55.
- Goodin, Robert E. “What Is So Special about Our Fellow Countrymen?” *Ethics* 98, no. 4 (July 1988): 663–86.
- Jackson, Frank. “Decision-Theoretic Consequentialism and the Nearest and Dearest Objection.” *Ethics* 101, no. 3 (April 1991): 461–82.

<sup>30</sup> I would like to thank Ben Lange, Roger Crisp, Serena Olsaretti, Areti Theofilopoulou, and audiences in Ghent and Belfast for their comments on earlier versions of this paper. I thank the Uehiro Foundation on Ethics and Education for its funding.

- Jeske, Diane. "Families, Friends, and Special Obligations." *Canadian Journal of Philosophy* 28, no. 4 (December 1998): 527–55.
- Keller, Simon. *Partiality*. Princeton: Princeton University Press, 2013.
- Kolodny, Niko. "Love as Valuing a Relationship." *Philosophical Review* 112, no. 2 (April 2003): 135–89.
- . "Which Relationships Justify Partiality? The Case of Parents and Children." *Philosophy and Public Affairs* 38, no. 1 (Winter 2010): 37–75.
- Lewandowski, Wojciech. "Parents, Special Obligations and Reproductive Genetics." In *The Ethics of Reproductive Genetics: Between Utility, Principles, and Virtues*, edited by Marta Soniewicka, 67–80. Dordrecht: Springer, 2018.
- Lillehammer, Hallvard. "Reproduction, Partiality and the Non-Identity Problem." In *Harming Future Persons*, edited by Melinda A. Roberts and David T. Wasserman, 231–48. Dordrecht: Springer, 2009.
- Olsaretti, Serena. "Liberal Equality and the Moral Status of Parent-Child Relationships." In *Oxford Studies in Political Philosophy*, vol. 3, edited by David Sobel, Peter Vallentyne, and Steven Wall, 58–83. Oxford: Oxford University Press, 2017.
- Petersen, Thomas Søbirk. "On the Partiality of Procreative Beneficence: A Critical Note." *Journal of Medical Ethics* 41, no. 9 (September 2015): 771–74.
- Pettit, Philip, and Robert Goodin. "The Possibility of Special Duties." *Canadian Journal of Philosophy* 16, no. 4 (December 1986): 651–76.
- Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13, no. 2 (Spring 1984): 134–71.
- Rakowski, Eric. *Equal Justice*. Oxford: Clarendon Press, 1991.
- Scheffler, Samuel. "Relationships and Responsibilities." *Philosophy and Public Affairs* 26, no. 3 (Summer 1997): 189–209.
- Steiner, Hillel, and Peter Vallentyne. "Libertarian Theories of Intergenerational Justice." In *Intergenerational Justice*, edited by Axel Gosseries and Lukas H. Meyer, 50–76. Oxford: Oxford University Press, 2009.
- Stroud, Sarah. "Permissible Partiality, Projects, and Plural Agency." In *Partiality and Impartiality: Morality, Special Relationships, and the Wider World*, edited by Brian Feltham and John Cottingham, 131–49. Oxford: Oxford University Press, 2010.
- Velleman, J. David. "Love as a Moral Emotion." *Ethics* 109, no. 2 (January 1999): 338–74.
- Williams, Bernard. "Persons, Character and Morality." In *Moral Luck: Philosophical Papers 1973–1980*, 1–19. Cambridge: Cambridge University Press, 1981.
- Wolf, Susan. "Morality and Partiality." *Philosophical Perspectives* 6 (1992): 243–59.

## PRIORITARIANISM A (PLURALIST) DEFENSE

*Shai Agmon and Matt Hitchens*

PRIORITARIANISM is the distributive view that welfare gains matter more, morally, the worse off you are.<sup>1</sup> A common and intuitively compelling objection to prioritarianism is that it wrongly treats cases involving one person (intra-personal cases) like cases involving more than one person (inter-personal cases), when they should be treated differently. In a nutshell, the objection goes as follows. A person is allowed, when faced with an intrapersonal choice between someone else's possible futures, to reason prudentially when choosing on their behalf.<sup>2</sup> She is not required to give special moral weight to the future in which the person for whom she is choosing would be worse off. However, when choosing how to distribute goods between multiple people, prudential reasoning on behalf of the group is not justified, as the claim of the person who is worse off should matter more (call this the *Moral Shift*). Thus, prioritarianism, which as an aggregative, impersonal view is committed to treating intrapersonal and interpersonal trade-offs similarly, cannot explain the Moral Shift.

Opinion is divided among philosophers as to whether this objection is powerful enough to reject prioritarianism.<sup>3</sup> In a 2009 article, Michael Otsuka

- 1 Throughout this article, for clarity and consistency, we refer to "welfare" rather than "utility" or "well-being." Our definition is not intended to differ in any meaningful way from other standard canonical formulations of prioritarianism. For examples, see Arneson, "Luck Egalitarianism and Prioritarianism," 340; Broome, *Weighing Goods*, 199; Parfit, "Another Defence of the Priority View," 401; O'Neill, "Priority, Preference and Value," 335; Segall, "In Defense of Priority (and Equality)," 344; Otsuka and Voorhoeve, "Why It Matters That Some Are Worse Off than Others," 176.
- 2 In all the cases to come, we assume that whatever set of relations is required for prudential concern (e.g., identity, psychological continuity, the prudential unity relations) holds, and holds to the same degree, between all the current and future people under discussion.
- 3 For example, McCarthy thinks this objection is sufficient to reject prioritarianism, but McKerlie and Rabinowicz do not. Interestingly, McCarthy, *contra* Otsuka and Voorhoeve, argues that prioritarianism should be rejected in favor of utilitarianism (rather than a pluralist egalitarianism), which does not recognize the Moral Shift. On McCarthy's view, the

and Alex Voorhoeve level a famous argument in favor of this proposition.<sup>4</sup> They claim that prioritarianism's inability to explain the Moral Shift means it ignores the unity of the individual, and hence should be rejected altogether. In this article, we show that their argument has a significant weakness, as its only answer to a standard prioritarian response is self-defeating. To explain briefly: a natural way in which the (pluralist) prioritarian can explain the Moral Shift is to appeal to the autonomy of the person being chosen for, as the prioritarian is choosing as the person would choose for herself. In order to preempt this, Otsuka and Voorhoeve assume the decision maker is a morally motivated stranger who knows (almost) nothing about the preferences and attitudes of the person in question. This assumption, we argue, makes the examples on which their argument relies incoherent, unless they rely on assumptions that most prioritarians would (we believe) find implausible. Thus, we argue that most prioritarians can both continue to appeal to autonomy to explain the Moral Shift and maintain their prioritarian commitments.

The article is structured as follows. First, we sketch the commitments of the prioritarian, and set out the objection to her view that we will be discussing. Next, we spell out the pluralist prioritarian's natural response to the Moral Shift—an appeal to the autonomy of the person being chosen for. We show that Otsuka and Voorhoeve can only head off this response, and so continue to press their objection, on the basis of two controversial assumptions: that welfare consists in a list of objective goods, and—if one takes an unorthodox but plausible view of risk aversion—that there is only a narrow range of rational risk aversions (or, analogously, that the prioritarian's weightings are extreme). These assumptions, we argue, significantly limit the scope of Otsuka and Voorhoeve's argument, and make it applicable for only a limited range of prioritarians. Along the way, we address possible responses on Otsuka and Voorhoeve's behalf, and show that they are unsuccessful.

## 1. PRIORITARIANISM

As mentioned, prioritarianism is the distributive view that welfare gains matter

---

Moral Shift is not justified to begin with. In this article, we assume that the Moral Shift exists and needs justification, and try to explain it in prioritarian terms. We do not address the claim that there should be no Moral Shift whatsoever. See McKerlie, "Dimensions of Equality"; Rabinowicz, "Prioritarianism for Prospects" and "Prioritarianism and Uncertainty"; McCarthy, "Utilitarianism and Prioritarianism II."

4 Otsuka and Voorhoeve, "Why It Matters That Some Are Worse Off than Others." Otsuka has defended and expanded on this critique in Otsuka, "Prioritarianism and the Separateness of Persons" and "Prioritarianism and the Measure of Utility."

more, morally, the worse off you are; the prioritarian unit of moral value is priority-weighted welfare. Prioritarianism can take a variety of forms, depending on how steeply the moral weighting applied to a unit of welfare decreases as the person receiving the welfare becomes better off. One limiting case is utilitarianism, which is prioritarianism with a weighting of one (i.e., no matter how badly off you are, an additional unit of welfare matters equally). The other is something like a maximin principle, which is prioritarianism with an infinite weighting (i.e., no welfare gains to the better off, no matter how large they are, could justify making the worse-off worse off, even by a minute amount).<sup>5</sup> We assume no one holds the view that welfare matters more if it goes to the better-off. Most prioritarians take an intermediate position, arguing that sufficiently large gains to the better-off can outweigh losses to the worse-off.<sup>6</sup>

## 2. OTSUKA AND VOORHOEVE'S ARGUMENT

Otsuka and Voorhoeve object to the priority view. Drawing on an example of Thomas Nagel's, they ask us to imagine two cases, structurally similar to the following cases.<sup>7</sup>

*Two-Person with Certainty:* There are two siblings, one who is disabled and one who is able-bodied. The disabled child would benefit from moving to the city, as he would have access to medical support; the able-bodied child would benefit from moving to the country, as she would be able to ride bikes and go on long walks.<sup>8</sup> The welfare gained by the able-bodied child in moving to the country is greater than the welfare gained by the disabled child in moving to the city. You, as a morally motivated stranger, must decide what to do for the children (they must live together). This choice is represented in Table 1.<sup>9</sup>

5 Sen, "Rawls versus Bentham," 302.

6 Arneson, "Luck Egalitarianism and Prioritarianism"; Parfit, "Another Defence of the Priority View"; O'Neill, "Priority, Preference and Value," 332–34; Segall, "In Defense of Priority (and Equality)."

7 Nagel, *Mortal Questions*, 123–24.

8 We use children in our examples throughout this article for consistency. We recognize that children are sometimes used in the literature as examples of agents without the capacity for autonomous choice, but we assume throughout that children do have this capacity, except where it is clearly indicated (by reference to the "Absence of Subjective Information Condition," which we introduce below).

9 Cf. Otsuka and Voorhoeve, "Why It Matters That Some Are Worse Off than Others," 187.

Table 1

|                          | Units of welfare |                   |
|--------------------------|------------------|-------------------|
|                          | Disabled child   | Able-bodied child |
| Send them to the city    | 40               | 50                |
| Send them to the country | 20               | 90                |

*One-Person with Risk:* There is one child with a particular diagnosis—she is 50 percent likely to be disabled, 50 percent likely to be able-bodied. If she is disabled, she would benefit from moving to the city, as she would have access to medical support; if she is able-bodied, she would benefit from moving to the country, as she would be able to ride bikes and go on long walks. The welfare gained by the child if she is able-bodied in moving to the country (over the able-bodied child living in the city) is greater than the welfare gained by the child if she is disabled in moving to the city (over living in the country). You, as a morally motivated stranger, must decide what to do for the child before you find out whether she will be disabled or able-bodied.<sup>10</sup> This choice is represented in Table 2.<sup>11</sup>

Table 2

|                         | Units of welfare   |                       |
|-------------------------|--------------------|-----------------------|
|                         | If she is disabled | If she is able-bodied |
| Send her to the city    | 40                 | 50                    |
| Send her to the country | 20                 | 90                    |

Otsuka and Voorhoeve’s argument relates to the Moral Shift between *Two-Person with Certainty* and *One-Person with Risk*. The prioritarian makes no distinction between these two cases. In both, sending the child(ren) to the country results in a higher total expected welfare than the other; but sending her/them to the city results in a higher total priority-weighted welfare.<sup>12</sup> Therefore, the prioritarian chooses the city in both cases, and does so for exactly the same strength of reason.<sup>13</sup> Otsuka and Voorhoeve, on the other hand, argue that the cases should

10 Structurally similar cases can be constructed with different units of welfare in which the same will be true for any prioritarian set of weightings.

11 Otsuka and Voorhoeve, “Why It Matters That Some Are Worse Off than Others,” 188.

12 Otsuka and Voorhoeve, “Why It Matters That Some Are Worse Off than Others,” 179–82, 188.

13 We are here considering the standard interpretation of prioritarianism, on which it is actual, not expected, welfare that receives priority weighting. For consideration of a view that attaches weight to both actual and expected welfare, see Parfit, “Another Defence of the Priority View”; Otsuka, “Prioritarianism and the Separateness of Persons,” 368, 375–80.

be treated differently. This is because the outcomes in *Two-Person with Certainty* are possible futures of two different children, whereas the outcomes in *One-Person with Risk* are both possible futures of the same child. This is a morally relevant fact for Otsuka and Voorhoeve. They believe that it is permissible to reason prudentially on behalf of the child, and “trade off” the expected welfare of her different futures, without attending to the distribution of the expected welfare between those different possible futures. This is due to the unity of the individual: it matters, they claim, that two possible futures are part of the same person’s future.<sup>14</sup>

To motivate this view, they ask us to imagine that the one-child asked, having been moved to the country and then having turned out to be disabled, “Why did you choose to send me to the country?” The morally motivated stranger could reply, “to maximize your expected welfare.” This response is not available in *Two-Person with Certainty*. In that case, if the children were moved to the country and the disabled child asked why this decision had been taken, the morally motivated stranger would have to reply, “to make someone, who was already better off than you, even better off.” Prioritarianism is insensitive to the presence of the additional prudential justification, and thus, the argument goes, insensitive to the intuitive force of the Moral Shift. Otsuka and Voorhoeve believe that you have more reason to send the one child to the country than the two children.<sup>15</sup> They contend that prioritarianism’s insensitivity to the prudential justification (i.e., to the Moral Shift) ignores the unity of the individual, and is therefore a serious moral shortcoming.

### 3. THE PRIORITARIAN RESPONSE: AN APPEAL TO AUTONOMY

However, if this were all there was to the example, the prioritarian would have an easy way out. Prioritarians need not be purists: they can be (and, we would argue, should be) pluralists, holding that more than one thing matters morally.<sup>16</sup> A defensible pluralist version of prioritarianism values both priority-weighted welfare and autonomy.<sup>17</sup>

As Gerald Dworkin notes, the concept of “autonomy” is often used by differ-

14 Otsuka and Voorhoeve, “Why It Matters That Some Are Worse Off than Others,” 180.

15 And, going further, Otsuka argues that you can simply reason prudentially on behalf of the child, without giving any prioritarian weight to benefiting her if she turns out to be disabled. See Otsuka, “Prioritarianism and the Measure of Utility,” 5.

16 O’Neill, “Priority, Preference and Value,” 343–44.

17 Otsuka and Voorhoeve acknowledge this option; see “Why It Matters That Some Are Worse Off than Others,” 186.

ent people to mean different things.<sup>18</sup> It is therefore worth spending some time defining what we mean by autonomy, as it forms a central part of the argument to come. Otsuka and Voorhoeve define autonomy for the purposes of their discussion as “a deference to [one’s] wishes regarding choices [one] has a right to make.”<sup>19</sup> There are at least two ways in which “deference to her wishes” could be taken. For this article, we distinguish between two kinds of autonomy, each of which is a possible interpretation of Otsuka and Voorhoeve’s definition.<sup>20</sup> One is deference to actual (self-regarding) preferences—we call this “superficial autonomy.”<sup>21</sup> Superficial here does not imply meaningless: superficial autonomy can really matter. Respecting actual preferences, even when they conflict with one’s underlying values or reasons, is respecting what Joseph Raz calls the *independence principle*. Raz suggests a range of cases for which it is applicable.<sup>22</sup> For example, if someone chooses a particular romantic partner, even if I know for a fact that her interest in that partner rests on false beliefs, and I further know that another partner best aligns with her core values, if I am in a position to choose her partner for her it seems that I should choose the partner she herself chose out of respect for her superficial autonomy.

The second kind of autonomy is what we will call “deep autonomy”—deference to one’s actual (self-regarding) values, expressed in the form of one’s ideal preferences. Ideal preferences are here understood as those that would have been autonomously formed on the basis of one’s deeply held values by an individual under conditions of full information, deliberation, and rationality (this differs from Otsuka and Voorhoeve’s conception of “ideal preferences,” as will become clear). For example, if I know that you are committed to living a healthy lifestyle, but mistakenly think that eating sugary cereal is the best way to achieve that goal, I am respecting your deep autonomy if I serve you oatmeal rather than Coco Pops for breakfast. This is because what I choose for you is what you would have chosen in ideal deliberative conditions, even if it is not what you would have actually chosen—I am deferring to your ideal preferences rather than your actu-

18 Dworkin, *The Theory and Practice of Autonomy*, 6.

19 Otsuka and Voorhoeve, “Why It Matters That Some Are Worse Off than Others,” 186.

20 Our distinction builds on a standard view in the literature on personal autonomy, according to which (roughly speaking) an agent acts autonomously when she acts as she does due to some mental state of her own regarding the matter. See Frankfurt, *The Importance of What We Care About*, 80–94, and *Necessity, Volition and Love*, 155–80; Shoemaker, “Caring, Identification, and Agency”; Jaworska, “Caring and Full Moral Standing”; Watson, “Free Agency.”

21 This is what Arneson calls “the simplest formulation of subjectivism”; see Arneson, “Autonomy and Preference Formation,” 42. We limit the discussion to self-regarding preferences and values for clarity and to align with Otsuka and Voorhoeve’s approach.

22 Raz, “The Problem of Authority,” 1014.

al preferences. This notion is rightly called “autonomy” because it still involves respecting what the individual themselves would choose: each of our value sets are different, and so we still need to know about someone’s values, or reasons, before we can choose on their behalf on the basis of deep autonomy.

Respecting both kinds of autonomy involves deferring to someone’s preferences, whether actual or ideal, and therefore—given the assumption that both kinds of preferences are self-regarding—reasoning prudentially on their behalf in one-person cases. Respecting superficial autonomy means reasoning prudentially from actual preferences, and respecting deep autonomy means reasoning prudentially from ideal preferences. It is for this reason that the pluralist prioritarian has a straightforward response to Otsuka and Voorhoeve’s example, as originally formulated: we have reason to do that which maximizes welfare in one-person cases because this is what respecting autonomy demands. In other words, the Moral Shift can be explained because autonomy can be appealed to in the one-person case, but not in the two-person case. The pluralist prioritarian could therefore agree with Otsuka and Voorhoeve that, all things considered, the child should be sent to the country, but could reject the contention that this means they must give up their prioritarianism, as it is compatible—due to their pluralism—with such an all-things-considered conclusion.

One might worry that this way around the argument commits prioritarians to always choosing for the person what they would choose for themselves in one-person cases. This would mean autonomy must be a dominant consideration for them in such cases. But this is not the case, as Otsuka and Voorhoeve recognize.<sup>23</sup> They acknowledge the fact that, even if the prioritarian denies that, all things considered, one should send the child to the country in *One-Person with Risk*, she must still contend with the fact that some shift in moral weighting is required when moving from two-person to one-person cases—there must be some Moral Shift. Even if it is not decisive, the fact that this prudential justification exists in one-person cases, where it does not in two-person cases, must make some difference to the moral calculus. And this can be accounted for by the pluralist prioritarian through an appeal to autonomy. Though autonomy need not always be decisive, it must always have some impact on deliberation. To the extent that there is a Moral Shift when moving from two-person to one-person cases, that weighting can always be accounted for by an appeal to autonomy, when one’s prioritarianism is pluralist. It is this subtler version of the pluralist defense that can save the prioritarian from the objection outlined above.<sup>24</sup>

23 Otsuka and Voorhoeve, “Why It Matters That Some Are Worse Off than Others,” 179.

24 For similar discussions regarding a subtler version of pluralist egalitarianism, see Parfit, “Another Defence of the Priority View,” 399; Temkin, “Egalitarianism Defended,” 780.

## 4. THE "ABSENCE OF SUBJECTIVE INFORMATION" CONDITION

Otsuka and Voorhoeve foresee the response of appealing to autonomy, and want to isolate considerations of priority-weighted welfare from this (legitimate) pluralist position. So they assume that the child is too young to have well-informed and rational preferences regarding the outcomes.<sup>25</sup> They believe that this condition, which we will call the *Absence of Subjective Information Condition* (ASIC), rules out considerations of autonomy. They do not clearly set the condition out as a complete absence of subjective information, but we believe this is the most plausible way to interpret it: if we only have completely unreliable and ill-informed subjective preferences (whether actual or ideal), then we effectively have no subjective information with which we can deliberate. Indeed, this is precisely the point of the condition, otherwise it would not isolate prioritarian considerations from considerations of (superficial or deep) autonomy. With this condition in place, there are no preferences (whether actual or ideal) to appeal to, and so, in this modified case, "there is ... no rival autonomy-based justification" available to the prioritarian for moving to the country.<sup>26</sup> Otsuka and Voorhoeve believe that, with this condition, they can continue to appeal to the notion of the expected welfare of the child with the ASIC in place. This can be done, they argue, with reference to what they call the child's "ideally rational and self-interested preferences."<sup>27</sup> These are not ideal preferences in the sense discussed thus far, i.e., preferences that would have been formed under ideal deliberative conditions, but something else: they idealize away not only misinformation, irrationality, and the like, but also any difference in underlying values, such that we can know what a person's "ideally rational and self-interested preferences" are without any subjective information. Since the preferences are completely idealized/abstracted in this way, they do not rely on any information about the actual child, and so the notion of expected welfare that depends on them can survive the introduction of the ASIC in a way that the pluralist prioritarian's appeal to autonomy cannot. In this way, then, Otsuka and Voorhoeve believe that they can sidestep the pluralist prioritarian defense, and press the case that prioritarianism ignores the unity of the individual.

We agree that the ASIC is necessary to rule out the pluralist prioritarian's ap-

25 Otsuka and Voorhoeve, "Why It Matters That Some Are Worse Off than Others," 187. For babies, or anyone completely without rational capacities, we can still identify universally pleasurable or painful goods, but these are basic, few in number, and do not significantly affect our argument; see note 29.

26 Otsuka and Voorhoeve, "Why It Matters That Some Are Worse Off than Others," 188.

27 Otsuka, "Prioritarianism and the Measure of Utility," 15.

peal to autonomy. However, it has further implications, which mean that Otsuka and Voorhoeve's argument causes no trouble for the majority of prioritarrians: the very move that blocks the appeal to autonomy also renders their argument, in most cases, toothless. This is because there is only one way of determining welfare for the outcomes with the ASIC in place, and that is through an objective-list account of welfare.

#### 4.1. *The ASIC Implies an "Objective-List" Account of Welfare*

Three standard types of accounts of welfare appear in the literature: hedonistic, preference-satisfaction, and objective-list accounts.<sup>28</sup> To sketch an outline of these accounts: hedonistic, or mental state, accounts hold that a person's welfare consists in their balance of pleasure over pain. Preference-satisfaction, or desire, accounts hold that a person's welfare consists in the extent to which a person's preferences are fulfilled. Finally, objective-list accounts hold that a person's welfare consists in the extent to which they possess "objective goods." These goods are such that, even if the person does not derive pleasure from them and would prefer not to have them (i.e., does not benefit from them subjectively), their welfare is improved through their possession thereof. These accounts are called "objective-list" accounts because proponents of such a view owe us, and often provide, a list of the goods that make us better off, regardless of our subjective views on the matter. So, for example, someone who is given a good education, even if they are completely indifferent to it or would have preferred not to be given it, can be said to be made better off by having the good of education on these accounts.

Each of these accounts has been met with well-known objections, and it is not the purpose of this article to evaluate their independent merits. Instead, we want to draw attention to a (rather obvious) feature of hedonistic and preference-satisfaction accounts that is not shared by objective-list accounts: they are dependent on the person's subjective information. If we know nothing about someone's preferences, values, psychological characteristics, and dispositions (i.e., if the ASIC is in place), it is normally impossible to determine whether a certain state of the world will increase or decrease their welfare, on a hedonistic account.<sup>29</sup> For preference-satisfaction accounts, the case is even clearer: without

28 Kymlicka, *Contemporary Political Philosophy*, 13–19; Parfit, *Reasons and Persons*, 493. For a contemporary modified version of this classification, see Dorsey, "The Hedonist's Dilemma." For recent criticism of the standard classification of welfare accounts, see Fletcher, "A Fresh Start for the Objective-List Theory of Well-Being"; Woodard, "Classifying Theories of Welfare." For convenience, we use the standard classification—the differences between the classifications do not affect our argument.

29 There are two exceptions to this rule, neither of which poses a threat to our argument as a

any knowledge of what someone's preferences are, we cannot know what would satisfy them. For both these accounts, then, if we are deprived of all subjective information about the person in question, we simply cannot say what impact some change would have on their welfare.<sup>30</sup>

It is worth noting that the connection between autonomy and subjective information is of a different character for the two subjective kinds of accounts of welfare. For preference-satisfaction accounts, the relationship is identical: what satisfies someone's preferences (whether actual or ideal) and what they would choose for themselves are one and the same. For hedonistic accounts, the relationship is more contingent: what will give someone pleasant mental states and what they would choose for themselves normally overlap. What is more, in actual cases, there is a strong epistemic connection, as the only way to understand what effect something would have on someone's mental states is normally to infer it from information about their preferences and values. In everyday life, we assess the welfare value of outcomes by comparing them with other possible outcomes: in action, by choosing an outcome (and therefore not choosing the other possible outcomes), and in speech, by expressing our enjoyment of an outcome (and therefore not expressing our enjoyment of the other possible outcomes).<sup>31</sup> For this reason, we tend to infer that a friend would prefer the park over the cinema when we know that they would enjoy the park (and we know

---

whole. The first is for those hedonists who believe that pleasure can be measured completely objectively, for example by measuring the levels of dopamine in the subject's bloodstream. For these hedonists, we accept that ASIC does not prevent them from ranking outcomes by their welfare-improving properties, but we take this view to be implausible and rarely held. Further, these hedonists face the epistemic problem of determining what effects going to the city or the country will have on the child's dopamine levels if they do not use subjective information as a guide. The second is for universally pleasurable goods (though it is not obvious that anything falls into this category). These goods can be distributed by the hedonist without any subjective information, but the small number of goods, if any, in this category puts strict limits on the importance of this fact. For the purposes of this essay, we treat dopamine-type views as special cases of objective-list accounts, with just one good on the list—the implausibility of this view means its inclusion in objective-list accounts does not greatly improve their attraction. And we ignore the distribution of universally pleasurable goods, since the goods under discussion in all the examples in the literature fall outside this category.

- 30 The argument to come also holds true for hybrid and pluralist accounts of welfare, which we address in note 33. For examples of hybrid accounts of well-being, see Raz, "The Role of Well-Being"; Parfit, *Reasons and Persons*, 501–2; Kagan, "Well-Being as Enjoying the Good," 253–55. For an example of a pluralist account, see Fletcher, "A Fresh Start for the Objective-List Theory of Well-Being."
- 31 Often when we say we enjoy something we actually mean that we would prefer it, given a certain set of choices in front of us.

nothing, let us assume, about how much they would enjoy the cinema)—and, if asked, we would say that we are choosing for them as we think, according to the available information, they would choose for themselves.

There are theoretical exceptions to this practical rule: most notably, Guy Fletcher's example of the ascetic, who prefers not to have pleasure because it would anger their deity, shows that hedonistic welfare and preferences can come apart in some cases (universally pleasurable goods might be another).<sup>32</sup> But for the most part, because preferences and hedonistic welfare substantially overlap, and because hedonistic judgments rely on preference information in normal cases, the connection between hedonistic welfare and autonomy is strong enough for the purposes of this article. Because this connection holds contingently and epistemically, it is more precise to say that, for the preference-satisfaction theorist, the ASIC rules out subjective information and therefore autonomy, whereas for the hedonist it rules out the basis for both determining welfare and determining autonomy. But in both cases, the presence or absence of subjective information tracks the ability of the prioritarian to appeal to autonomy, whether necessarily or contingently, and therefore in both cases the ASIC undermines Otsuka and Voorhoeve's examples. This is not true of objective-list accounts. If education improves someone's welfare, then, regardless of their subjective experience or view of that education, they are made better off by it. The objectivity of such accounts allows their proponents to tell people, without reference to their subjective information, that something has made them better off. This is a special property of objective-list accounts: only they can tell us about someone's welfare in completely abstract conditions, when we are deprived of all subjective information about that someone. For the same reason, it is only with an objective-list account of welfare that Otsuka and Voorhoeve's notion of "ideally rational and self-interested preferences," which in turn is only required because of the ASIC, makes sense. It is a conception of preference entirely abstracted from facts about the individual who holds the preference, such that every person in a situation has the same such preference, and so it requires an objective-list account of welfare to have any content at all.

Consider again *One-Person with Risk*. With the ASIC in place, neither hedonistic nor preference-satisfaction accounts of welfare can make sense of this example, because the example relies on determinate units of welfare that could not be generated on the basis of available information. Without knowing anything about the child's attitudes toward riding bikes and going on long walks, how could the morally motivated stranger know that the country will give her a welfare of 90 if she is able-bodied, on either kind of account? On hedonistic

32 Fletcher, "A Fresh Start for the Objective-List Theory of Well-Being," 211–14.

accounts, there is no information about what will give her pleasurable mental states; on preference-satisfaction accounts, there is no information about her preferences. She could hate the outdoors and enjoy access to museums (which the city provides) much more; or love the outdoors far more than 90 suggests. Because there is no way to construct such a numerical measure of the child's welfare, when the ASIC is in effect, a proponent of either of the subjective accounts of welfare must reject the example as incoherent and need not respond to it.<sup>33</sup>

Objective-list theorists, however, are pressed by the example. For them, it is entirely coherent to say that a world in which someone has a certain set of opportunities for outdoor activity is better for their welfare than a world in which they would have a certain other set of health states (depending, of course, on the exact constitution and weighting of their objective lists). If the child were to turn around, *post facto*, and say, "I would have liked the city much better," the objective-list theorist has the resources to respond, "Your welfare, as expressed by the satisfaction of your 'ideally rational and self-interest preferences,' is higher here, regardless of what you think or feel." For prioritarians who subscribe to an objective-list account of welfare, then, Otsuka and Voorhoeve's objection remains a telling one. So appealing to autonomy is not enough to reject Otsuka and Voorhoeve's argument altogether, but it is enough to severely curtail its scope.

#### 4.2. Possible Responses: Ideal Preferences with the ASIC?

If Otsuka and Voorhoeve were to respond to our argument successfully, while avoiding assuming an objective-list account of welfare, they would need to show that there is some way of deriving their "ideally rational and self-interested preferences" without knowing anything about the actual preferences or values of the child, and without relying solely on objective facts about the choices. In their original article, Otsuka and Voorhoeve make an attempt in this direction, which we argue is unsuccessful. They try to ground the quantification of the expected welfare of the child in aggregated hypothetical preferences expressed by other people. As they write, in relation to a similar example:

Surveys indicate that people who imagine themselves in such a predica-

33 In the special case in which there is no subjective information about the child, hybrid accounts of welfare—which include both subjective and objective elements and therefore for which subjective information is fundamental—imply that there is insufficient information to decide which option is best for the child, and thus Otsuka and Voorhoeve's objection to them is unsuccessful. Pluralist accounts of welfare (essentially objective-list accounts in which desire satisfaction is one good) imply that there is sufficient information, as subjective information is not fundamental to their account—it is just one good on the list. Therefore, we should treat the pluralist-account theorist as a variant on the objective-list theorist, and address the same objections to her.

ment would typically be indifferent between (i) receiving the treatment that might raise one from a state of very severe impairment to this state of severe impairment and (ii) receiving the treatment that might entirely cure one of the slight impairment.<sup>34</sup>

These impairments were fully described to those responding to the surveys, and—using the “standard gamble” method—indifference points were reached for the relevant changes in welfare. Otsuka and Voorhoeve rely on this fact in the construction of their later examples (on which ours are based), in order to fill out what they mean by the welfare of the various people in question. It might be thought that this process grounds the welfare of the child in such a way as to avoid the argument above, as there is no need to appeal to an objective-list account. By resting on the aggregate hypothetical preferences of other people, this view might be thought to avoid relying on the actual preferences of the child, but be subjectively determined and thus compatible with hedonistic or preference-satisfaction accounts—and so meet the objection outlined above.

There are two ways to interpret this attempt, and the distinction between what is evidence of welfare and what constitutes welfare is useful to understand them. Something can be used as evidence of welfare when it indicates that someone’s welfare has gone up, regardless of the underlying philosophical account of what welfare consists in. Mainstream economics often uses preference satisfaction in this way; the rate at which people smile might also be thought of a measure of welfare. Something is constitutive of welfare, on the other hand, when welfare consists in it—the standard debate about welfare, sketched above, is a debate about what constitutes welfare. Whether the survey data is understood as evidence of welfare or constitutive of welfare, the argument bottoms out as a poor way of understanding what would make the child better off. On the one hand, Otsuka and Voorhoeve might be using the hypothetical preferences of others as a way of approximating what the child would want: as a measure of their welfare. They might think that the fact that a representative group of people claims to be indifferent between two health states is good evidence for another person also being indifferent between those states. But this is just an appeal to autonomy, as preferences remain the constitution of well-being: it is an attempt to approximate the (actual/ideal) preferences of the child with reference to the (hypothetical) preferences of representative others—and an imperfect one, at that. If this is the way in which they seek to cash out this response, then they reopen the door for prioritariness to appeal to autonomy (and end up with a much

34 Otsuka and Voorhoeve, “Why It Matters That Some Are Worse Off than Others,” 172.

worse way of getting at the child's preferences than if they were to drop the ASIC), and therefore neutralize the response's purpose.

On the other hand, they might be arguing that the hypothetical preferences of others constitute welfare. But this is very implausible: it is hard to see why it would be the case that the aggregated views of others play a role in constituting what will make a different individual better off. This is an account of welfare that contradicts all three standard accounts. If I am a hedonist or a preference-satisfaction theorist, then what matters is how the states of health accord with the individual: the individual might have very different preferences from the group surveyed, so relying on their views will not help. But if I am an objective-list theorist, the survey results will not matter either (unless perhaps only objective-list theorists were surveyed): the situation itself tells me everything I need to know, and how people believe they would respond to it is neither here nor there. So falling back on this defense would imply an even stronger version of our original argument: that Otsuka and Voorhoeve's example has nothing to say about anyone who has taken a view on the "account of welfare" question. Deriving welfare from the hypothetical preferences of others therefore offers no support to Otsuka and Voorhoeve, whichever way it is interpreted.

#### 4.3. *Possible Responses: A Weaker ASIC?*

Another possible response to our argument has been suggested by Otsuka in correspondence. Rather than assuming an objective-list account of welfare, he suggests the following assumption could be made:

We know facts about [the child's] psychological dispositions and character traits, including their general level of risk appetite. On the basis of this, we're able to idealize in a manner that yields determinate idealized preferences over gambles. We don't, however, know whether they would (in their non-idealized state) actually have these preferences if confronted with the particular choices among gambles. Maybe they would. Maybe they wouldn't.

In essence, this assumption is a weaker ASIC: rather than assuming away all subjective information, we simply assume away information about actual preferences, leaving enough information to determine what the child's ideal preferences (in this article's sense) would be. Determining ideal preferences in this way might be thought to make the examples coherent, while avoiding a reliance on an objective-list account of welfare.

This response is unsuccessful because the weaker version of the ASIC that it implies no longer performs the function required: it does not successfully rule

out considerations of autonomy. Although only knowing about the child's psychological dispositions and character traits, rather than their actual preferences, prevents the prioritarian from appealing to superficial autonomy, this weaker ASIC gives her precisely the information required to appeal to deep autonomy. With knowledge of the child's underlying values, assuming we are ideal deliberators, we can determine what they would choose if they were ideal deliberators (i.e., determine what their ideal preferences are). As we have said, both superficial and deep autonomy could matter. In section 3, we set out some cases in which superficial and deep autonomy come apart, and—following Raz—suggested that sometimes one should take precedence, and sometimes the other. But if we only have enough information to determine what respecting one kind of autonomy demands, then it is clear that someone who respects autonomy in general ought to respect that kind of autonomy. As an instance of this, in the one-child case, if we have enough information to determine what respecting the child's deep autonomy demands, then the pluralist prioritarian has a reason to choose in line with this: a reason directly derived from their respect for autonomy. This weaker ASIC therefore does not perform the task Otsuka and Voorhoeve demand of it, and so a stronger assumption—of an objective-list account of welfare—remains necessary to render their argument coherent. In fact, because both the determinate ideal preferences of the child and considerations of autonomy rely on the same subjective information, one will always track the other: if we have enough information to avoid assuming an objective-list account of welfare, we will have enough information to appeal to autonomy, and vice versa. However the strength of the ASIC is varied, then, it will not enable Otsuka and Voorhoeve to avoid this dilemma.

With these possible responses met, the conclusion of our argument remains undefeated. Without information about the subjective information of the child, there is no way to level the objection Otsuka and Voorhoeve describe without relying on an objective-list account of well-being to fill out their notion of “ideally rational and self-interested preferences.” If we are objective-list prioritarrians, the objection remains as compelling as before. If we are not, were we to be faced with a case like *One-Person with Risk*, we would have to respond that more information is required, that without knowing what the preferences of the child are we simply could not say what would maximize the priority-weighted utility of the child. And if more information were to be provided, then appealing to autonomy would be back on the table.

## 5. THE ASIC IMPLIES A RANGE OF RATIONAL RISK ATTITUDES

Section 4 narrowed the scope of Otsuka and Voorhoeve's argument to those prioritarrians who subscribe to an objective-list account of well-being; this section will argue that even among these prioritarrians, if we accept an unorthodox but plausible account of the normative relationship between risk aversion and welfare, most will have little to worry about.

"Orthodox" decision theory offers one way to represent an agent's preferences over prospects, according to which the value that a person assigns to an act can be determined by the conjunction of two functions: one that assigns utilities to its possible outcomes (a "utility function"), and one that assigns probabilities to those outcomes, conditional on each act being chosen (a "probability function"). On this view, we can construct cardinal rankings of outcomes (an ordering that puts numerical values on the utilities in each outcome) by referring to these two functions only. This idea is often formalized using the expected utility framework of von Neumann-Morgenstern.<sup>35</sup> To cardinalize preferences, in this framework, gambles are set up between the various outcomes. The gambles take the form, "You have an  $X$  percent chance of  $A$  or a  $Y$  percent chance of  $B$ . Which do you prefer?" With a sufficient number of such gambles, for various values of  $X$  and  $Y$ , the ranking can be cardinalized using the information resulting from the choices.<sup>36</sup>

Choosing an option in these gambles requires, as an input, our risk aversion—if we have a very low chance of  $A$ , we might be more drawn to choosing  $B$  because we are averse to the risk of leaving empty-handed. So we cannot go through the cardinalizing process sketched above without relying on facts about our attitudes to risk. On the orthodox view, we can only come to numerical values for the utilities of outcomes if we have included risk aversion in the cardinalizing process. Risk aversion is not distinguished from any other desire attitudes. It is treated as any other preference, or, more accurately, it is taken into account in the process of the preference formation.<sup>37</sup> To be risk-averse with respect to expected utilities is to be irrational, on the orthodox view. If a proponent of Otsuka and Voorhoeve's argument holds some version of this view, we have nothing further to say in relation to risk aversion specifically—it is to be treated as just another kind of preference, and so the arguments in section 4 apply.

However, one of the major critiques of the orthodox view is that it fails to

35 Von Neumann and Morgenstern, *Theory of Games and Economic Behavior*.

36 For an elaboration on the orthodox view, see Stefánsson and Bradley, "What Is Risk Aversion?" 2; Buchak, *Risk and Rationality*, 1–2, 10–36.

37 Stefánsson and Bradley, "What Is Risk Aversion?" 2.

distinguish between preferences over goods and preferences over risk itself: risk aversion is treated as just another preference, albeit a general one that applies to all gambles.<sup>38</sup> Many argue that because of this failure, the orthodoxy does not capture the phenomenology of risk attitudes.<sup>39</sup> “Unorthodox views” in choice theory focus on this issue, and argue that risk aversion should be considered separately. Lara Buchak has suggested that a third parameter be added to the mix of the orthodox recipe: a separate function that measures the agent’s attitude to risk.<sup>40</sup> As a result, she proposes a “risk-weighted expected utility” theory, in which—because a risk function is included in the analysis of what matters—it is not irrational to be risk-averse with respect to expected utilities.<sup>41</sup>

We refrain from providing arguments for either of these views. Nevertheless, the distinction between orthodox and unorthodox views of risk aversion is relevant to our current endeavor, as on unorthodox views such as Buchak’s—which we believe have merit—Otsuka and Voorhoeve are forced to make a further assumption in order for their examples to work. This assumption narrows the scope of the applicability of their argument even further, as this section will argue.

On the unorthodox view, roughly speaking, in everyday life, if we are faced with a risky prospect, we take into account, in addition to our preferences for the possible outcomes, our general attitudes to risk. For example, when it comes to deciding whether to insure one’s house, two people might derive exactly the same utilities from the outcomes (equally happy if the house remains in good shape, equally annoyed if it burns down), and have the same understanding of the probability of a fire, but if the first is more risk-averse than the second, she might buy an insurance policy that the second thinks is not worth the money. Neither person would normally be condemned as irrational, because risk attitudes, like preferences, are seen as subjective. People tend to think that a variety of levels of risk aversion (defined here as how much negative weight is given to variance in outcomes) can be rational.

This fact is relevant to Otsuka and Voorhoeve’s argument. Even granting that we are objective-list prioritarrians, the ASIC means that we do not have access to any subjective information about the child, including their risk attitudes. To

38 Stefánsson and Bradley, “What Is Risk Aversion?” 2.

39 Watkins, “Towards a Unified Decision Theory”; Hansson, “Risk Aversion as a Problem of Conjoint Measurement.”

40 Buchak, *Risk and Rationality*, 49–50.

41 Other defenses of the view that it is rationally permissible to be risk averse with respect to one’s own good (whether in the form of one’s values or one’s hedonistic welfare) include Broome, *Weighing Goods*; Okasha, “On the Interpretation of Decision Theory.”

complete their examples, on the unorthodox view, they must tell us something about the risk aversion of the child—as it turns out, the only plausible assumption they could make regarding risk aversion further reduces the number of objective-list prioritarians who must respond to their argument.

We might have a very risk-averse child, Avril; we might have a child, India, who is indifferent to risk; or (less likely) we might have a daredevil, Darren, who is risk-seeking. The ASIC means Otsuka and Voorhoeve need to make an assumption about what a rational risk attitude looks like, otherwise the decision about the child's future will be lacking a vitally important piece of information. Since the examples are set up with the welfare in each of the outcomes cardinalized, the assumption needs to spell out what weightings it is rational to apply to expected welfare.

To illustrate the difference between Avril, India, and Darren, we can consider what they would choose in the following gamble.

Table 3

|          | Units of welfare |       |
|----------|------------------|-------|
|          | Heads            | Tails |
| Choice A | 1,000            | 0     |
| Choice B | 501              | 499   |

The expected value of both A and B is 500. For India, this is all there is to the gamble, so India would be indifferent between the two choices. For Avril, the expected value of B is given extra weight, because the variance between the outcomes is lower and the worst possible outcome (ending up empty-handed) is avoided. So Avril would choose B. Darren, finally, would be excited by the thrill of the gamble of A, and so would choose A.

So, what assumption about the rationality of risk attitudes do Otsuka and Voorhoeve have to make? For their examples to work, on the unorthodox view, they must assume that there is a range of rational risk attitudes: anyone whose level of risk aversion falls within this range is acting rationally, but anyone who is extremely risk-averse (or risk-seeking) is not. Indeed, Otsuka implies that he would agree with this assumption.<sup>42</sup> Not only does this assumption allow the examples to work, it also chimes with common sense: someone who goes to extreme lengths to insure themselves against bad outcomes, paying thousands of pounds a month in premiums, would normally be considered irrational, fearing bad outcomes too much and being irrationally inattentive to the cost of avoiding

42 See Otsuka, "Prioritarianism and the Measure of Utility," 8, where he refers (in another context) to a risk aversion that is "low, yet not so low as to be irrationally so."

them. With this reasonable assumption, Otsuka and Voorhoeve's examples are complete: we are able to account for the fact that the child might be Avril, India, or even Darren (if any risk-seeking attitudes fall within the rational range).

This necessary assumption has consequences, however, that further weaken the overall argument against prioritarianism. First, since it introduces the notion of a rational range, there will be cases in which either the city or the country is in the rational self-interest of the child, and so both options will be permissible. How many such cases there will be depends on how widely the range is drawn, a matter that need not concern us here. In cases in which the expected welfare outcomes are sufficiently close, rational self-interest will be indeterminate. On reflection, this is how things should be: there are many choices in everyday life, like the prospect of moderately expensive house insurance, for which we might think rational self-interest is not decisive. So this does not pose a problem for Otsuka and Voorhoeve—though it means that, in such cases, their objection to the prioritarian is much weakened, as the prioritarian choice (choice B, for instance) is a permissible choice.

More troubling for their argument is the implication that, if a prioritarian's weightings are sufficiently moderate, they will always choose permissibly. Prioritarian weightings are a close analogue of risk-attitude weightings, as both are weightings applied to welfare to generate final value decisions, which (when the attitude is one of risk aversion) add weight to choices that avoid the worst outcomes. Risk aversion manifests, in cases like these, as a direct tendency to reduce the variance between outcomes, and prioritarianism, by adding more moral weight to the worse outcomes than to the better outcomes, does the same. Indeed, this analogy has been used by Derek Parfit to defend the rationality of prioritarianism.<sup>43</sup> If the prioritarian weightings are, by analogy, "within" the range of rational risk attitudes, then the prioritarian choice will always be a rational one, and Otsuka and Voorhoeve's objection to prioritarianism—that in some cases it chooses that which is not in the ideally rational self-interest of the child—will no longer hold true.

To see this, consider again *One-Person with Risk*:

Table 2

|                         | Units of welfare   |                       |
|-------------------------|--------------------|-----------------------|
|                         | If she is disabled | If she is able-bodied |
| Send her to the city    | 40                 | 50                    |
| Send her to the country | 20                 | 90                    |

43 Parfit, "Another Defence of the Priority View," 423.

The expected value of sending her to the city is 45, and the expected value of sending her to the country is 55: this is what makes the example useful in Otsuka and Voorhoeve's argument. But the variance of sending her to the city is only 10, whereas the variance of sending her to the country is 70. If someone is risk-averse (within the rational range) they will therefore add weight to the option of going to the city, and this could mean, in this case, that the city is actually a more attractive prospect than the country. In a strongly analogous way, the prioritarian will add more weight to the lower-value outcomes, and less to the higher-value outcomes, and come to the same conclusion. Importantly, one could describe the prioritarian weightings and the risk-averse weightings using variants of the same formula: one would be over outcomes, and another would be over variances in outcomes, but each level of prioritarianism would track one and only one level of risk aversion. This means that, if the risk aversion is within the rational range, then so is the prioritarianism, and therefore the prioritarian will always choose a permissible option. Of course, in this example sending her to the country would also be rational (assuming that an indifference to risk is also within the rational range of risk attitudes), but this would be only one of the rational, or permissible, choices.

What does this imply, on the unorthodox view of risk aversion? It means that any moderate prioritarian who applies non-extreme weightings to welfare will always choose permissibly—by Otsuka and Voorhoeve's criterion of acting in the rational self-interest of the child—in these one-person cases. Of course, the notion of "moderate" is tied to the definition of the rational range of risk attitudes. If the range is small, then few prioritarians can be sure of always choosing permissibly. But as long as such a range is admitted, there will be a class of prioritarians who can safely ignore Otsuka and Voorhoeve's arguments.

Otsuka would dispute this conclusion (and in so doing, give an indication that he might be sympathetic to the unorthodox view of risk aversion). He has argued that risk aversion does not necessarily tell in favor of prioritarianism, so even if someone were being risk-averse when choosing for another they would sometimes choose the utilitarian over the prioritarian option. To demonstrate this, he gives the example of coming across a stranger in the wilderness who has just suffered an accident. He has a 50 percent chance of being disabled, and a 50 percent chance of full health. You have the option of what he calls the "risky treatment," which would make the stranger's condition less severe if he turns out to be disabled, but would leave him slightly disabled if he were to be in full health. To complete the symmetry with *One-Person with Risk*, the difference in welfare between "slightly disabled" and "full health" is greater than the difference between "disabled" and "less severely disabled." In this case, Otsuka claims, the

risk-averse person would not administer the “risky treatment.” This is because they will not be “*taking* any risk on his behalf if you leave him alone.”<sup>44</sup>

We believe that Otsuka’s understanding of risk aversion as relying on the doing versus allowing distinction is misguided. As his emphasis shows, and as the wilderness example reveals more generally, for Otsuka the treatment is risky only insofar as you must actively choose it, whereas you merely allow the accident to take its course. However, when risk aversion is understood, as outlined above, as an aversion to variance in outcomes, the wilderness example does not tell against its identification with prioritarianism. Administering the treatment makes the least bad outcome less bad (at a cost to the value of the best outcome), and thereby reduces the variance in outcomes. When the distinction is drawn in this way, it is clear that the morally motivated stranger, truly choosing—with risk aversion—on behalf of the unconscious person, would choose the (not risky) treatment. It is possible that, independently, the additional worry about doing versus allowing would give them a separate reason to avoid the treatment; but it would not be risk aversion, properly understood, that would give them such a reason. If we are risk-averse in the sense of aversion to variance in outcomes, then a range of rational risk aversions will always make room for prioritarians, and the identification between the two views will remain consistent.

Otsuka and Voorhoeve might argue that, though the moderate prioritarian always chooses a permissible option, they do so for the wrong reason: they continue to wrongly ignore the unity of the individual, and treat the two possible futures of the person as distinct outcomes, when they should be considered as part of one life. This objection is unaffected by our argument in this section, but, in our view, the sting is taken out of it when it does not come with any impermissible implications. Further, prioritarians tend to subscribe to the view that one can coherently talk about something being good *sub specie aeternitatis* (this is often part of what it means to be a prioritarian), so an objection that calls this into question without providing further worrying counterexamples would be an external, rather than an internal, criticism.

## 6. CONCLUSION

These two objections, each an implication of the ASIC, work together to mean that Otsuka and Voorhoeve’s argument, and by extension the general objection to prioritarianism, has no force against the majority of pluralist prioritarians. Those who are vulnerable to the objection, and so must answer the charge of wrongly ignoring the unity of the individual, are only those who subscribe to

44 Otsuka, “Prioritarianism and the Measure of Utility,” 9.

both an objective-list account of welfare and either: (a) orthodox decision theory, or (b) an unorthodox decision theory in which the range of rational risk aversions is narrow (or, analogously, one coupled with an extreme set of prioritarian weightings). For these prioritarians, and there are likely to be some, the arguments of this article are no help: they continue to owe us a reason why the Moral Shift is not reflected in their moral calculus. For all others, the ASIC does not make sense, and so appealing to autonomy remains a decisive justification for the Moral Shift—their prioritarianism thus remains undefeated.<sup>45</sup>

*shaiagmon@gmail.com*  
*matthitchens@live.co.uk*

#### REFERENCES

- Arneson, Richard J. "Autonomy and Preference Formation." In *In Harm's Way: Essays in Honor of Joel Feinberg*, edited by Jules L. Coleman and Allen Buchanan, 42–75. Cambridge: Cambridge University Press, 1994.
- . "Luck Egalitarianism and Prioritarianism." *Ethics* 110, no. 2 (January 2000): 339–49.
- Broome, John. *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Blackwell, 1991.
- Buchak, Lara. *Risk and Rationality*. Oxford: Oxford University Press, 2013.
- Dorsey, Dale. "The Hedonist's Dilemma." *Journal of Moral Philosophy* 8, no. 2 (2011): 173–96.
- Dworkin, Gerald. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.
- Fletcher, Guy. "A Fresh Start for the Objective-List Theory of Well-Being." *Utilitas* 25, no. 2 (June 2013): 206–20.
- Frankfurt, Harry G. *The Importance of What We Care About*. Cambridge: Cambridge University Press, 1988.
- . *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999.
- Hansson, Bengt. "Risk Aversion as a Problem of Conjoint Measurement." In *Decision, Probability, and Utility*, edited by Peter Gärdenfors and Nihl-Eric Sahlin, 136–58. Cambridge: Cambridge University Press, 1988.

<sup>45</sup> We would like to thank Campbell Brown, Susanne Burri, David Enoch, Ittay Nissan-Rozen, Mike Otsuka, Shlomi Segall, and Alex Voorhoeve for helpful comments on earlier versions of this article.

- Jaworska, Agnieszka. "Caring and Full Moral Standing." *Ethics* 117, no. 3 (April 2007): 460–97.
- Kagan, Shelly. "Well-Being as Enjoying the Good." *Philosophical Perspectives* 23 (2009): 253–72.
- Kymlicka, Will. *Contemporary Political Philosophy: An Introduction*. 2nd ed. Oxford: Oxford University Press, 2002.
- McCarthy, David. "Utilitarianism and Prioritarianism II." *Economics and Philosophy* 24, no. 1 (March 2008): 1–33.
- McKerlie, Dennis. "Dimensions of Equality." *Utilitas* 13, no. 3 (November 2001): 263–88.
- Nagel, Thomas. *Mortal Questions*. Cambridge: Cambridge University Press, 2012.
- Okasha, Samir. "On the Interpretation of Decision Theory." *Economics and Philosophy* 32, no. 3 (November 2016): 409–33.
- O'Neill, Martin. "Priority, Preference and Value." *Utilitas* 24, no. 3 (September 2012): 332–48.
- Otsuka, Michael. "Prioritarianism and the Measure of Utility." *Journal of Political Philosophy* 23, no. 1 (March 2015): 1–22.
- . "Prioritarianism and the Separateness of Persons." *Utilitas* 24, no. 3 (September 2012): 365–80.
- Otsuka, Michael, and Alex Voorhoeve. "Why It Matters That Some Are Worse Off than Others: An Argument against the Priority View." *Philosophy and Public Affairs* 37, no. 2 (Spring 2009): 171–99.
- Parfit, Derek. "Another Defence of the Priority View." *Utilitas* 24, no. 3 (September 2012): 399–440.
- . *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Rabinowicz, Wlodek. "Prioritarianism and Uncertainty: On the Interpersonal Addition Theorem and the Priority View." In *Exploring Practical Philosophy: From Action to Values*, edited by Dan Egonsson, Jonas Josefsson, Bjorn Petersson, and Toni Rønnow-Rasmussen, 139–65. Burlington: Ashgate, 2001.
- . "Prioritarianism for Prospects." *Utilitas* 14, no. 1 (March 2002): 2–21.
- Raz, Joseph. "The Problem of Authority: Revisiting the Service Conception." *Minnesota Law Review* 90, no. 4 (2006): 1003–44.
- . "The Role of Well-Being." *Philosophical Perspectives* 18 (2004): 269–94.
- Segall, Shlomi. "In Defense of Priority (and Equality)." *Politics, Philosophy and Economics* 14, no. 4 (November 2015): 343–64.
- Sen, Amartya. "Rawls versus Bentham: An Axiomatic Examination of the Pure Distribution Problem." *Theory and Decision* 4, no. 3–4 (February 1974): 301–9.

- Shoemaker, David W. "Caring, Identification, and Agency." *Ethics* 114, no. 1 (October 2003): 88–118.
- Stefánsson, H. Orri, and Richard Bradley. "What Is Risk Aversion?" *British Journal for the Philosophy of Science* (forthcoming).
- Temkin, Larry S. "Egalitarianism Defended." *Ethics* 113, no. 4 (July 2003): 764–82.
- Von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. 6th ed. Princeton: Princeton University Press, 2004.
- Watkins, John W. N. "Towards a Unified Decision Theory: A Non-Bayesian Approach." In *Foundational Problems in the Special Sciences: Part Two*, edited by Robert E. Butts and Jaakko Hintikka, 345–79. Dordrecht: D. Reidel, 1977.
- Watson, Gary. "Free Agency." In *Agency and Answerability: Selected Essays*, 13–33. Oxford: Clarendon Press, 2004.
- Woodard, Christopher. "Classifying Theories of Welfare." *Philosophical Studies* 165, no. 3 (September 2013): 787–803.

## REASONS: WRONG, RIGHT, NORMATIVE, FUNDAMENTAL

*Errol Lord and Kurt Sylvan*

ACCORDING to *reasons fundamentalism*, all normative properties are analyzable in terms of reasons.<sup>1</sup> Famously, some of the analyses offered by reasons fundamentalists face the *wrong kind of reasons problem*. This problem first appeared in the literature on the buck-passing account of value, which says in its simplest form that what it is for something to be valuable is for there to be sufficient reasons to have a pro-attitude toward it.<sup>2</sup> This simple view fails, many worry, because there can be reasons for having pro-attitudes toward things that have nothing to do with their value. Contrasting cases like *Beauty* and *Extra Credit* provide an illustration:

*Beauty*: Jane is a first-year graduate student in art history. She has loved art all her life, but is just now getting the opportunity to see Europe's masterpieces through her graduate program. She sees the *Mona Lisa* in person for the first time. She is enthralled by its symmetry, depth, and enigmatic tone.

*Extra Credit*: Jack is a smart aleck on a high school trip to Paris. He could not care less about art and finds the Louvre to be a total bore. Predictably, he has been causing all sorts of trouble. As they enter the room with the *Mona Lisa*, Jack starts photobombing tourists' shots of the painting. In a fit of frustration, Jack's high school history teacher says to him, "Will you

- 1 Prominent reasons fundamentalists include Scanlon, *Being Realistic about Reasons*; Parfit, *On What Matters*; Schroeder, *Slaves of the Passions*; Skorupski, *The Domain of Reasons*; and Dancy, *Ethics without Principles*. We are here understanding reasons fundamentalism as a claim about the internal structure of the normative—i.e., as the claim that reasons are the fundamental items *within the normative domain*. It is consistent with reasons fundamentalism of the sort considered here that reasons might turn out to be naturalistically grounded. While "reasons fundamentalism" could be used in a stronger sense that implies nonnaturalism, that is not how we are using it.
- 2 See Rabinowicz and Rønnow-Rasmussen, "The Strike of the Demon"; D'Arms and Jacobson, "The Moralistic Fallacy" and "Sentiment and Value."

just be enthralled by *something*? I will give you extra credit if you show some appreciation and respect for the *Mona Lisa*.”

Jane responds to some reasons to be enthralled by the *Mona Lisa*. Intuitively, her reasons are of the right kind. The painting’s depth, symmetry, and enigmatic tone are part of what make it valuable. Since Jane’s enthralment is a response to these reasons, it is plausibly *fitting*. Jack’s reasons are different. The fact that Jack will get extra credit if he is enthralled does not help to explain why the *Mona Lisa* is valuable. Jack’s teacher could have provided a similar reason for him to be enthralled by something with no value—a saucer of mud, for example. If Jack were to be enthralled just to get extra credit, it is plausible that his enthralment would be unfitting. This is because he is enthralled for the wrong kind of reason.

It seems that buck-passers about value need to understand value solely in terms of the right kind of reasons.<sup>3</sup> Otherwise it seems their view badly overgeneralizes. A sufficient incentive to have a pro-attitude would entail that the object of that pro-attitude is valuable. This is a bad prediction. So buck-passers face pressure to explain the distinction between the right kind of reasons (RKRS) and the wrong kind of reasons (WKRS).

While the literature on buck-passing is the most familiar context in which the wrong kind of reasons problem arises, it is easy to see that the same problem arises for other reasons fundamentalist analyses. To see why, one need only consider how reasons fundamentalists might go about analyzing other normative properties that can be picked out with words of the form “ $\phi$ -ible/-able” and “ $\phi$ -worthy,” like the properties of being *credible*, *trustworthy*, *despicable*, *noteworthy*, and so on. In each case, it is plausible that there can be sufficient reasons for  $\phi$ -ing an  $X$  that do not bear on  $X$ ’s  $\phi$ -ability. For this reason, it is clear that the wrong kind of reasons problem is a much more general problem.

Interestingly, this point shows not only that reasons fundamentalists have many WKR problems, but also that *many theorists* have WKR problems. For many theorists can and should find it attractive to analyze *some* things in terms of reasons. Indeed, as Schroeder emphasizes, it is attractive to think that a reasons-based account has to be the right account of properties like credibility, despicability, trustworthiness, noteworthiness, and so on, whatever one thinks about the broader reasons fundamentalist program.<sup>4</sup> Reflections on this fact have led some—most notably, Schroeder—to suggest that fundamentalists are

3 We add “it seems” because some buck-passers (e.g., Skorupski, *The Domain of Reasons*; Parfit, *On What Matters*) deny that there are wrong-kind reasons, and hence appeal to reasons, period, rather than right-kind reasons. But this response neglects the generality of the phenomenon, as we will see.

4 Schroeder, “Value and the Right Kind of Reason.”

off the hook.<sup>5</sup> The fact that their account would badly overgeneralize without a principled distinction between *RKRS* and *WKR*s does not undermine their approach, they insist, because the *WKR* problem is everybody's problem, and the challenge for fundamentalists will be answered if the more general question has an answer. And so they conclude that they face no special challenge, thanks to the generality of the problem.

Our main goal in this paper is to argue that this conclusion is mistaken and indeed gets things backward. We think that reflection on the alleged generality of the distinction between right- and wrong-kind reasons suggests that life is *harder*, not easier, for reasons fundamentalists. With this goal in mind, we defend two main claims. The first claim is:

1. The apparent generality of the *WKR* problem suggests that the class of right-kind reasons is wider than the class of normative reasons; the reasons that determine "fittingness" are not essentially normative in any sense stronger than that associated with any arbitrary standard of correctness. This gives rise to what we call the *Right Kind of Reasons Problem*.

The second claim is:

2. If claim 1 is true, then some special problems—including a potentially fatal dilemma—arise for fundamentalists.

After defending claim 2, we will evaluate several possible escape routes for fundamentalists.

Our plan is as follows. In section 1, we begin by considering the case for the generality of the *WKR* problem at greater length. We argue that any adequate solution to the more general *WKR* problem will imply that not all *RKRS* are normative reasons. Thus, we cannot determine which considerations properly figure into reasons-based analyses simply by solving the *WKR* problem. We also need to figure out which *RKRS* are normative reasons, and why they are normative reasons. This is the *Right Kind of Reasons Problem*. To solve the *RKR* problem, one needs to solve two other problems, which we call the *Normativity Problem* and the *Rationale Problem*. In section 2, we argue that reflection on these problems raises a dilemma, which we call the *Fundamentalist's Dilemma*. In section 3, we canvass three responses to this dilemma. The first two are reductive naturalist views—the first, a Humean view, and the second, a novel constitutivist view. The third is a new form of quietism, one that enjoys some advantages over other quietist views. We suggest that the constitutivist option is the best. We leave it to future work to decide whether this is a strike against reasons fundamentalism

5 Schroeder, "Value and the Right Kind of Reason."

or an argument for constitutivism. In section 4, we summarize the key morals and conclude.

## 1. THE RIGHT KIND OF REASONS PROBLEM

### 1.1. *The Generality of the Distinction and Why Not All RKRS Are Normative Reasons*

If the WKR problem is everybody's problem, there must be a distinction between RKRS and WKRS that all theorists can and should acknowledge. Thankfully for reasons fundamentalists, there is a plausible candidate for such a distinction, and it is a *very* general distinction. As Schroeder points out, wherever one finds a *standard of correctness* governing an activity, one finds a distinction among considerations recommending potential moves in the activity that could naturally be dubbed a distinction between reasons of the right kind and reasons of the wrong kind.<sup>6</sup> In particular, for any standard of correctness, there are certain considerations that bear on whether an act would help one conform to the standard and others that are simply irrelevant. Insofar, then, as one is engaged in an activity governed by a standard of correctness, there are certain considerations to which one ought to be sensitive relative to the standard, and others to which one ought not be sensitive relative to the standard.

It is natural enough to describe this distinction as a distinction between reasons of the right kind and reasons of the wrong kind. To appreciate this fact, one need only consider examples like the following pairs of cases:

#### *Rope Pair*

*Good Boy Scout:* Kenny is a precocious and studious Boy Scout. He has learned how to tie most of the knots in the Boy Scouts handbook. He is currently trying to tie a half hitch. Placing the left portion of the rope over the right would be an efficient step toward producing a half hitch as described by the book. Kenny chooses to manipulate the rope accordingly with this fact in mind.

*Bad Boy Scout:* Billy is a terrible Boy Scout. He has it out for Kenny, and likes to mess with Kenny's sense of Boy Scout decency. So when he sees Kenny practicing his knot-tying skills, he decides to have some fun with him. He decides to offer him \$20 to deviate from the book when tying his half hitch. Kenny sees the utility of the \$20 and is thus disturbed, just as

6 This point is central to both Schroeder's (see especially "Value and the Right Kind of Reason") and D'Arms and Jacobson's (see especially "Sentiment and Value") approaches to the WKR problem.

Billy intended. Kenny gives in this one time and moves the rope in a way that will at best lead to a very bad example of a half hitch.

*Table Pair*

*Proper Butler:* Mr. Carson is a proper English butler. He always sets the table to the standards of the proper English tradition. Arranging the forks to the left of the knives is one way to help get the table in order. He chooses to arrange the silverware accordingly with this fact in mind.

*Devious Valet:* Thomas is a devious valet. He does not care much for the traditions except when he feels they are to his advantage. He decides one day that it is to his advantage to incentivize breaking tradition for Mr. Carson. So he makes it so that the only way for Mr. Carson to avoid embarrassment in front of Lady Mary is by breaking with table-setting tradition and putting the knives to the left of the forks.

In Good Boy Scout, Kenny's choice is appropriate relative to the standard of correctness established by the handbook's description of a half hitch (though he has not at the end of the example yet produced a half hitch).<sup>7</sup> Why? The natural explanation is that Kenny's choice properly takes into account how various available ways of manipulating the rope would make a difference to whether he correctly produces a half hitch. Considerations that bear on whether a certain way of manipulating the rope would correctly produce a half hitch are reasons of the right kind relative to the standard. Hence, Kenny's choice is appropriate relative to the standard because it is sensitive to reasons of the right kind relative to the standard. In Bad Boy Scout, Kenny is not doing what he should be doing by the lights of the standard. He is responding to an irrelevant consideration relative to the standard by choosing to move the rope in a way that will merely produce some money for him.

Similar things can be said about the Table Pair. In Proper Butler, Mr. Carson's choice is appropriate by the lights of the relevant standards. A natural ex-

7 Note that we here assume that a choice or act can be *appropriate* relative to a standard of correctness even if it does not constitute *conformity* to the standard. We will also use the word "fitting" in this way, though we realize that some use "fitting" as a synonym for "correct." We assume that appropriateness can be naturally enough analyzed in terms of reasons of the right kind relative to the standard, where these are understood as considerations that bear on whether acting in the relevant way would make a contribution toward conformity. We do not similarly assume that the standard of correctness is analyzable in terms of reasons; indeed, we will ultimately suggest that as long as standards of correctness are not inherently normative, it is permissible for a reasons fundamentalist to explain the right-kind/wrong-kind reason distinction by appeal to them.

planation is that he chooses for reasons that are relevant and good relative to the standard: he chooses to move the silverware in a way that will help him to set the whole table in the proper English way. In *Devious Valet*, Mr. Carson is led off the proper path by considerations that are irrelevant to the norms governing such arrangements.

These cases draw attention to a distinction that seems worth calling a distinction between reasons of the right kind and reasons of the wrong kind. Notice that this distinction is not just a distinction among motivating or explanatory reasons. Before Kenny decides to place the left portion over the right, there are various moves open to him. In choosing which move to make, there are facts to be borne in mind, like:

RR: Placing the left portion over the right is the most efficient way to complete the first step toward correctly tying a half hitch.

To those attempting a half hitch, RR recommends placing the left portion over the right at Kenny's stage. If it were causally possible to skip this step, RR might not conclusively recommend the move. Accordingly, RR's apparent force may vary depending on the causal possibilities. But assuming there is no way to skip the step, RR does conclusively recommend the move to the half hitcher. And it does so even if Kenny has not yet registered this fact, and hence is not among his motivating reasons.

To the extent, then, that it is plausible to call this fact a reason of the right kind, it seems not to be merely a motivating or explanatory reason. It is a consideration carrying a recommending weight that modulates the strength of the criticism that the half hitcher would face if he pursued a more cumbersome strategy. And the criticism need not be that he does not conform to the standard: he might eventually bumble into conformity. This fact suggests that the right-kind/wrong-kind distinction at issue is not just a distinction between whether or not one is doing what conforms to the standard. Each choice and move can be assessed for a kind of appropriateness or fittingness relative to the standard. This appropriateness is a function of how well one is responding to considerations relevant to the attainment of the standard.

Observations of this kind make it plausible not just that there is an RKR/WKR distinction that everyone can and should acknowledge, but also that there is a task worthy of everyone's attention: the task of explaining why certain considerations are relevant to this distinctive kind of criticism, while others are irrelevant despite having great, even conclusive, normative significance. Indeed, this task is compulsory for anyone interested in the grounds of the criticism. Again, it is plausible that whether a move would be the most appropriate move relative to the

standard is determined by whether it is the move most strongly recommended by these considerations, where some considerations can recommend with less than sufficient force (e.g., *this move would help a bit (though others are more helpful)*).

Of course, this task does not look daunting: the difference is clearly tied to the bearing of the considerations on whether one would meet the standards that constitutively govern the activity. Precisely this fact, one might hope, saves reasons fundamentalists from their problem. For, one might hope, their problem can then be solved as a special case of this easier and more general problem.

Appreciating this point, Schroeder uses the generality of the problem as inspiration for a particular account of what it is for something to be a right kind of reason.<sup>8</sup> According to this account, the right kind of reasons are always relativized to *activities*. They are the reasons that everyone engaged in the activity has *in virtue of* engaging in that activity. More officially:

*Activity:* What it is for  $r$  to be a right kind of reason to  $\phi$  relative to some activity  $a$  is for  $r$  to be a reason for anyone engaged in  $a$  to  $\phi$  in virtue of being engaged in  $a$ .

Activity is poised to make plausible extensional predictions. It does not seem like *anyone* tying knots has the reason provided by Billy—only Kenny does. But anyone tying Boy Scout knots does have the reasons provided by the book. The nature of Boy Scout rope tying seems to guarantee this. Similarly, not all the good butlers have the reason provided by Thomas—only Mr. Carson does. But all butlers have the reasons provided by tradition. The nature of butlering seems to guarantee this.

There are two basic thoughts behind Activity. First, wrong-kind reasons are not tied to the nature of the activities in the right kind of way. Second, there is something too local about wrong-kind reasons. Extensional predictions aside, we think that there is something very powerful behind these thoughts. The wrong kind of reasons do not essentially bear on whether one is helping to fulfill the aim of the activity, while the right kind of reasons do, and are considerations that everyone engaged in the activity ought to consider on pain of meriting a kind of criticism. Although we regard these thoughts as genuine insights, we will not assume that Activity in particular is true.<sup>9</sup> We merely assume in what follows

8 Schroeder, "Value and the Right Kind of Reason."

9 We will also not be assuming Schroeder's account of what it takes for  $r$  to be a reason for anyone engaged in  $a$  in virtue of being engaged in  $a$ . In fact, we think that Sharadin shows that Schroeder's view cannot work ("Schroeder on the Wrong Kind of Reasons Problem for Attitudes" and "Reasons Wrong and Right"). We, like Sharadin, think that this sort of agent neutrality should be explained in terms of the constitutive standards of the activity (although in personal communication Sharadin reports he now thinks appealing to con-

that some account that vindicates these insights has to be right, and Activity is at least a good stab in the right direction.

Before proceeding further, it is worth considering an objection to Activity in order to highlight what we will be assuming. The objection is that Activity puts the cart before the horse by insisting that what is fundamental are the activities; what is fundamental, according to the objector, are the reasons themselves. We can then derive activity-types from the reasons however we wish. For example, the objector might insist that what is fundamental to epistemic normativity are reasons for certain reactions. The activity of believing will then just be whatever activity purports to be sensitive to such reasons.

Reflection on standard buck-passing accounts of value encourages this objection. For such accounts appear to proceed on the assumption that what is fundamental are the reasons for various reactions. Once those reasons are picked out, we can then determine the activities (if we are so inclined). Activity denies that this is the order of explanation. Instead, it holds that the activities come before the (right-kind) reasons. The nature of the activities themselves determines the right-kind reasons.

While we recognize that this picture is natural—especially upon reflection on the basic structure of buck-passing accounts of value—we think that it ultimately misses the key lessons of the WKR problem. The fundamental lesson of the WKR problem is that *something* must be said in order to pick out the facts that are the reasons that serve in analyzing properties like value. This lesson of course does not force us to Activity. But once this point is appreciated, the virtues of Activity become apparent, for it does seem plausible that there is a *correlation* between certain activities and the right-kind reasons. When you couple this fact with the need to explain why right-kind reasons are right-kind, it starts to look appealing to think that the nature of the activities are what explain why the right-kind reasons are right-kind.

That said, we are not endorsing Activity. We are only endorsing four commitments that Activity helps to illustrate: first, that we need an explanation of what makes right-kind reasons right-kind; second, that one can draw the right-kind/wrong-kind distinction wherever there are standards of correctness; third, that the WKRS are wrong-kind in virtue of not being tied to the standards in the right way; fourth, and finally, that the RKRS are not local and idiosyncratic—rather,

---

stitutive standards is too constrictive). We suspect Schroeder does not go for this because he is afraid that this will be smuggling normative notions into his account of RKRS, which would in turn undermine his fundamentalism. We think that this is a mistake because it is a mistake to think that constitutive standards are genuinely normative. This will be spelled out in more detail below (see especially section 3.2).

they are reasons shared by those who are tied to the standards in some relevant way.<sup>10</sup> Since Activity illustrates the last three commitments so nicely, we will largely proceed as if it is true. This is merely for the sake of concreteness and simplicity. Assuming a concrete view will make it easier to state the problems at the heart of the paper.

Given these commitments, however, we must conclude that not all RKRS are normative reasons. For standards of correctness *per se* are normative only in an anemic sense. They set up some standard, but not all are genuinely normative. Explaining the distinction between anemic and genuine normativity is a hard task, but there are clear examples on either side: norms of etiquette and grammar, on the one hand, and epistemic norms, prudential norms, and moral norms, on the other.<sup>11</sup> All we assume here is that being a standard of correctness is not sufficient for being genuinely normative. This claim is one that even fundamentalists who think normativity comes fairly cheaply can embrace, for no fundamentalists would think there is necessarily a good reason to satisfy any standard whatsoever (though some believe in a vaster array of reasons than others). Yet it seems all standards generate a distinction between right-kind and wrong-kind reasons. So, for the standards that are not genuinely normative, we get RKRS that are not normative reasons.

To see this point vividly, consider cooking. There are many standards of correctness that purport to govern the cooking of *cacio e pepe* (a Roman pasta dish—literally, “cheese and pepper”).<sup>12</sup> For there are loads of recipes for making this dish. Many of them conflict. For any given standard corresponding to a recipe, there will be RKRS to do certain things—e.g., to add extra cheese table-side. But—and here is the important point—not all of these RKRS are normative reasons, even for those who are engaged in the cooking of *cacio e pepe*. As any self-respecting Roman will tell you, many of these recipes lead to inferior *cacio e pepe*. The RKRS associated with those standards plausibly have no more genu-

10 We are fine with one reading our conclusion in this section conditionally. If one accepts these commitments, then it is plausible that not all right-kind reasons are normative reasons. Establishing this conditional would still be an important result given that prominent views accept the commitments.

11 Perhaps the most famous invocation of the difference between genuine normativity and anemic normativity is Foot, “Morality as a System of Hypothetical Imperatives.” For further discussion, see Wodak, “Mere Formalities”; Parfit, *On What Matters*; Copp, “Moral Naturalism and Three Grades of Normativity”; Woods, “The Authority of Formality”; Baker, “Skepticism about Ought Simpliciter”; Plunkett and Shapiro, “Law, Morality, and Everything Else”; and McPherson, “Against Quietist Normative Realism.”

12 In Errol Lord’s opinion, it is the most delightful thing one can put in one’s mouth when it is prepared well.

inely normative significance than other anemic standards, such as the norms of etiquette and grammar.

It does not follow from this point that the right-kind/wrong-kind distinction that all should acknowledge is not a distinction in *reasons*, as we will again emphasize in section 1.4. There is a criticism to which one is open in ignoring considerations relevant to satisfying the standards of correctness governing an activity. And these considerations have weights linked to their bearing on how well one would be satisfying the standards that govern the activities, in the attributive-value sense of “well.” They hence have the functional profile of reasons: they are recommending considerations with weights whose balance determines a kind of criticism. But the genuine normative significance of both the criticism and the considerations that help to explain it is as much an open question as the normativity of criticisms of instrumental rationality and of the subjective reasons associated with requirements of instrumental rationality (which we also assume are indeed reasons, though perhaps not always normative reasons if Humeanism is false!).

Indeed, there is pressure for the reasons fundamentalist who wants to write off the WKR problem as everybody’s problem to insist that this distinction is a distinction among reasons. For the mere fact that a more general wrong-kind/right-kind distinction can be drawn does not show that there is a problem that everyone faces parallel to the WKR problem. A genuine analogue of the problem would arise only if there were a kind of assessment that ought to be understood as determined by the balance of considerations worth calling “reasons.” For only if there is independent reason for everyone to analyze some form of assessment in terms of reasons of the right kind would it be true that there is a WKR *problem* for everyone. It is not enough that there be a distinction that everyone can understand. For there might be a special obligation for certain theorists to explain this distinction owing to their need to invoke it in their theory, and only this fact generates a problem.

The upshot is that the generality of the WKR problem shows that not all RKRS are normative reasons, for some standards of correctness are normatively anemic. These standards generate RKRS that are not plausibly normative reasons. So, in order to determine which considerations are relevant for reasons-based analyses—in order to figure out what the normative reasons are—one has to determine which RKRS are normative reasons, and why they are normative reasons. This is the Right Kind of Reasons Problem. It has the same form as the WKR problem, for that too is a problem concerned with differentiating some favorers from others. It is widely assumed that once we sort the wrong-kind favorers from the right-kind favorers, we will know which favorers to appeal to in

reasons-based analyses. We have argued that this is false; we also need to know which right-kind favorers flow from robustly normative standards. To provide this story is to provide a solution to the RKR problem.

### 1.2. How to Solve the Right Kind of Reasons Problem

We agree that the WKR problem is everybody's problem. For similar reasons, we think that the RKR problem is everybody's problem.<sup>13</sup> In this subsection we will provide a schema for solving it.

Recall how we got stuck with the WKR problem. We start trying to analyze complex normative properties in terms of normative reasons. We then see that there are some facts that seem to recommend reactions in the wrong way. This observation leads us to search for an account of what it is for a fact to recommend a reaction in the right way. But as it turns out, a fact can recommend in the right way and yet *still* not be a normative reason. So, solving the WKR problem is not sufficient for being in a position to pick out the normative reasons that are supposed to be analyzing the complex stuff. This extra problem is the RKR problem.

To solve the RKR problem we need to carry out two extra steps. First, not only do we need an account like Activity, we also need an account about which activities are normatively relevant.<sup>14</sup> That is, we need to know which activities, if any, are such that the RKRS associated with those activities are necessarily normative reasons. Once we have an account of this, we will be in a position to determine which favorers can figure in reasons-based analyses. We call the problem of determining which activities are robustly normative the *Normativity Problem*. Solving this problem is the first step toward solving the RKR problem.

In order to solve the Normativity Problem, one must have a story about which activities are normatively relevant. Now, we recognize that there are some obvious candidates: believing, intending, and desiring, for example, will plausibly be on the list. The hard part is not coming up with a *list* of activities that are normatively relevant. So solving the Normativity Problem does not look that daunting. However, it is harder to give an account of what these activities have in common that makes them normatively relevant. We call this the *Rationale Problem*. Solving the Rationale Problem is the second step toward solving the RKR problem. With a solution to both the Normativity and the Rationale prob-

13 As we mentioned above, we are happy to weaken this to the claim that the RKR problem is a problem for everyone who accepts that the right view of RKRS has the four features we elucidate in section 1.1.

14 Again, we will speak of activities because we find it illuminating. One need not appeal to activities *per se*. One could also appeal to standards of correctness or perhaps some other notion that plays the role that activities play in Activity.

lems, one will have an account of which activities are robustly normative and an explanation of why *those* are the robustly normative activities.

So far problems have only multiplied. We started off with the WKR problem. Now we have the RKR problem, which naturally breaks down into the Normativity Problem and the Rationale Problem. As we have said, the WKR and RKR problems are structurally on all fours. They are problems for everyone interested in using normative reasons to analyze normative phenomena. Further, fundamentalists face both problems in full generality. In other words, the success of the fundamentalist's entire metanormative theory hinges on solving both problems. Without a solution to both problems, the fundamentalist lacks the resources to even get going. This makes it all the more unfortunate that the fundamentalist has a particularly hard time solving the RKR problem. We will see this shortly. First, though, we will elucidate the RKR problem more by considering an analogous problem that only targets buck-passing about value. This will naturally lead to an attempt to dismiss the problem out of hand. After dismissing that dismissive response, we will explain why the fundamentalist has a particularly hard time solving the RKR problem.

### 1.3. Analogous Problems for Buck-Passers about Value *Simpliciter*

To further understand the RKR problem, it is helpful to compare it to an analogous but more local problem for a particular fundamentalist analysis—namely, the buck-passing account of value *simpliciter*.<sup>15</sup> In examining this more local problem, we will discover some points that help address a dismissive response to the RKR problem, which we rebut in the next subsection.

According to the simple buck-passing account of value *simpliciter*, what it is for *X* to be valuable *simpliciter* is for there to be sufficient reasons to have some pro-attitude toward *X*. This simple account, of course, faces the WKR problem. But as Schroeder points out, it is plausible that everyone should accept a buck-passing account of certain evaluative properties, like enviability and amusingness. He concludes that for this reason the buck-passer is off the hook.

Although considering properties like admirability and enviability points to a WKR/RKR distinction that everyone should want to explain, life is harder for buck-passers about value *simpliciter* for this very reason. The problem cases that arise for these accounts point to a notion of a right-kind reason and a correspond-

15 We do not assume that reasons fundamentalists are committed to the existence of value *simpliciter*, or that buck-passing accounts of value are exhausted by buck-passing accounts of value *simpliciter*. We focus on this case to explain the kind of problem we are raising, which could also be raised for buck-passing accounts of goodness-for (though perhaps not for accounts of attributive goodness, which is not clearly normative).

ing notion of fittingness unsuited for underpinning a buck-passing account of value *simpliciter*. This suggests that the real problem for buck-passers—that of explaining why certain reasons are relevant to the grounding of value *simpliciter* and others are not—remains even after they produce an adequate account of the WKR/RKR distinction. This is the localized version of the RKR problem.

To see this clearly, consider first the example of “amusing.” Close cousins to famous points made by Justin D’Arms and Daniel Jacobson bring out the challenge.<sup>16</sup> The key point is that the notion of a right-kind reason that makes the following analysis true is insensitive to goodness and badness *simpliciter*:

*Amusing*: What it is for *X* to be amusing is for there to be right-kind reasons to be amused by *X*.

The features of a joke or situation that make it amusing may well be features that make it bad *simpliciter*. There is something funny about someone repeatedly making a prudentially bad kind of mistake (this is the whole idea of slapstick comedy), and about absurd and incongruous instantiations of badness (imagine a house being struck by lightning one hundred times in a row). Fans of dark humor will find the absurd badness of the world funny in itself. On a natural first reading of the buck-passing account, this seems impossible. For the buck-passing account holds that the features we have reason to have pro-attitudes like amusement toward are good *simpliciter* (and in virtue of the fact that we have such reasons). This appears to ignore the role that the internal standards of funniness play. In particular, it seems to ignore that those standards seem insensitive to goodness and badness *simpliciter*. Those standards only care about humor, as it were. They embrace it wherever it is found, even in the bad.

One could try to deny that these things are amusing. But to do so would, as D’Arms and Jacobson stress, involve pointing to features that are wrong-kind reasons relative to the standards that spell out what it takes for something to be funny. That is, to think that these things are not funny simply because they are bad *simpliciter* is akin to thinking something is not funny because it is morally bad. It is to point to wrong-kind reasons relative to the standards governing funniness. Our point, again, is that the standards governing funniness are not directly sensitive to goodness and badness *simpliciter*. This is a cousin of D’Arms and Jacobson’s point that the standards of funniness are not directly sensitive to moral badness/impermissibility.<sup>17</sup>

16 See D’Arms and Jacobson, “The Moralistic Fallacy.”

17 It is worth spelling out the exact relationship our point has to D’Arms and Jacobson’s main point. Their main point is that it is a conceptual mistake to think that the fact that something is morally bad/impermissible is a reason not to be amused. We are making a parallel

Importantly, this is not a local point about amusement. It is plausible that emotions present their objects as meeting or failing to meet some standard of correctness, in some sense of “present.” But it is easy to imagine emotions that present their objects as meeting standards of correctness that fail to line up with anything good or bad *simpliciter*. More familiar emotions provide further examples. Consider emotions like *shock* at  $x$  and their corresponding correctness conditions (i.e.,  $x$ 's being *shocking*). Or if you do not like that example, consider religious emotions like *penitence*, which is fitting only if one has sinned. There are conditions under which these emotions are fitting, but from this nothing follows about whether anything of normative significance is encoded by these feelings.

It is easy to imagine other possible emotions that would illustrate the same moral. We might not have names in ordinary language for some of these emotions. But we may have other signs—consider an emoticon with someone wearing sunglasses giving a thumbs-up, which represents that to which it is responding as *cool*. One could imagine parallel emoticons for *fly*, *swag*, *rad*, etc. Supposing that we could coin words for these emotions of the form EMOT, we could imagine corresponding adjectives of the form EMOT-able. Right-kind reasons for EMOT-ing  $X$  will be evidence that  $X$  is EMOT-able. But the EMOT-ability conditions will be given by a standard of correctness that may not map onto anything of value or disvalue *simpliciter*.

What these reflections suggest is that we need a distinction between right-kind and wrong-kind considerations in favor of pro-attitudes and con-attitudes that flows from their built-in standards, without regard to whether these standards track value *simpliciter*. But if so, then the following revised version of the buck-passing account is false:

*Right-Kind Buck-Passing about Value Simpliciter*:  $X$  is valuable (/disvaluable) *simpliciter* iff there are sufficient right-kind reasons for having some pro-attitude (/con-attitude) toward  $X$ .

We can see that this cannot be right by thinking about cases like amusement, enviability, and shock. Amusingness, enviability, and shockingness do correlate

---

point against a reply the buck-passer might give to our suggestion that features that are bad *simpliciter* cannot be funny. To give that response is to ignore the fact that the standards that govern what is funny do not analytically rule out that features that are bad *simpliciter* are funny in virtue of being bad *simpliciter*. We are also making a point that goes beyond D'Arms and Jacobson's—namely that paying attention to the internal standards of amusement puts pressure on the buck-passer to explain which pro-attitudes count. This is because our point about amusement shows that amusement is not a pro-attitude that should ground goodness *simpliciter*. This illustrates the buck-passer's analogue to the Normativity Problem by showing that she has the task of telling us which pro-attitudes count.

with right-kind reasons for the attitudes of amusement, envy, and shock. But there is an open question about whether these properties are values/disvalues *simpliciter*. If so, we should reject Schroeder's case for getting buck-passers about value *simpliciter* off the hook. Perhaps everyone should expect there to be a solution to the WKR problem for "amusing" and "shocking." But it will not give the buck-passer everything she needs. It actually shows that she needs more than the general distinction between WKR/RKR for pro-attitudes. She also needs to tell us which pro-attitudes are relevant to the analysis of value *simpliciter*.

The buck-passer about value *simpliciter* now faces two problems. First, she must tell us which pro-attitudes are such that right-kind reasons for them ground value *simpliciter*. This is her analogue of the Normativity Problem. Second, the buck-passer must explain what these attitudes have in common, such that the presence of sufficient RKRS for bearing them to *X* entail that *X* is valuable. This is the buck-passer's analogue of the Rationale Problem.

Notice that the first problem differs from a superficially similar "problem" sometimes mentioned for buck-passers. It has, of course, always been a *question* for buck-passers which pro-attitudes are the ones to which they should at bottom appeal. Sometimes this question is regarded as raising a problem. But there is not a problem—just an interesting question—until it is shown that not all right-kind reasons for pro-attitudes are value-grounding in the robust sense of "value" that buck-passers were originally trying to analyze. But we can see from reflecting on the generality of the WKR problem that not all right-kind reasons for pro-attitudes are value-grounding in that sense. Hence, it is incumbent on buck-passers to solve their analogues of the Normativity and Rationale problems.<sup>18</sup>

#### 1.4. Dismissing a Dismissive Response

Reflection on the foregoing special case helps address a dismissive response to our problems. One might have imagined some reasons fundamentalists, such as Parfit or Scanlon, agreeing that there is a right-kind/wrong-kind distinction that arises relative to any arbitrary standard of correctness, but insisting that this distinction is not necessarily a distinction *among reasons*. The right-kind considerations that we are claiming are not normative reasons are, these theorists might

18 One option, of course, is to insist that the pro- and con-attitudes that play the relevant theoretical roles in the buck-passing account are primitive. This would sidestep the local versions of the Normativity and Rationale objections raised here. We have two things to say about this. First, this would not show that the problems are not problems for the buck-passer. Indeed, to accept this form of primitivism is a way of accepting that there is a problem. Second, primitivism is the option of last resort. We should investigate whether there are viable explanations before we become primitivists. We think that there are views that offer explanations. We explore some below.

claim, *not reasons at all*. So, these theorists might insist that they do not have to say which attitudes or activities are relevant. Call this the *dismissive response* to the RKR problem.

Before giving our main response to this reply, it is worth noting that this reply is not available to all fundamentalists. Primitivist fundamentalists like Parfit and Scanlon can make this move, but fundamentalists like Schroeder who are also Humeans cannot. Schroeder does think—unsurprisingly, given his Humeanism—that one can have reasons to fulfill the aim of some activity simply in virtue of being engaged in that activity. If one is trying to tie a knot, one presumably wants the knot to be tied and that desire just does, on a Humean view, generate certain reasons. (We will return to this below.) While Parfit and Scanlon might deride activity-relative reasons, such derision would be theory-driven and, we think, not terribly convincing.

More importantly, the Dismissive Reply is implausible as a response to the localized problem for the buck-passing account. Some reasons fundamentalists might antecedently have been fine denying that right-kind considerations in an activity governed by some standard of correctness are right-kind reasons in that activity. But it is not plausible to deny that the amusing properties of a joke are reasons to be amused by it. It makes good sense to be amused by things that are amusing, whatever else might be said about one's amusement. The properties of a joke can also provide more or less of a case for being amused, and hence play a contributory favoring role, as reasons do. We see no theory-neutral reason to deny the strong intuition that if a joke is amusing, there are reasons to be amused by it. Since features that are bad *simpliciter* are not necessarily right-kind reasons not to be amused (and some things may be amusing *because they are absurdly bad*), the funny-making features will not be relevantly defeated by other right-kind reasons. So, the point remains there: sufficient right-kind reasons for some pro-attitudes do not generate value *simpliciter*.

It is worth emphasizing that the point we are making here is a more local one that is intended to clarify the structure of our main point. Although it is structurally parallel to our main point, it is not the same point. What we are arguing here is that the existence of sufficient reasons of the right kind for having a pro-attitude toward *X* does not imply that *X* is valuable *simpliciter*. Even if one were to agree that the reasons in these cases are genuinely normative, our conclusion here would remain. It is unclear that being amusing entails being good *simpliciter* to any degree or in any way, or that being enviable entails being good *simpliciter* to any degree or in any way. We do not think it is plausible to claim that a deeply unequal world would be good in a way because it would entail that some people are enviable, where this is a way of being good. (For this reason, the analogue of

our Normativity Problem presently under discussion might be better called the *Value Problem*.)

Accordingly, the problem here is to explain which right-kind reasons for pro-attitudes are relevantly value-grounding, given that not all are. We think this problem is illuminatingly analogous to the problem of explaining which right-kind reasons in general are relevantly normativity-grounding. But the main difference between the problems simply reinforces the strength of the original Normativity and Rationale problems. For the assumptions needed to get these problems going are much more minimal. While there might be some shred of plausibility in insisting that all evaluative properties correspond to ways of being good *simpliciter*, there is no shred of plausibility in the thought that all standards of correctness correspond to ways of being genuinely normative. Yet for someone engaged in an activity, considerations that bear on whether acting in some way would contribute to satisfaction of the activity's constitutive standard of correctness have the same claim to being reasons as reasons for envy and amusement. Insensitivity to both bears on the appropriateness of certain forms of criticism, where the degree of appropriateness is modulated by how the relevant considerations balance out.

The bottom line, then, is that one should not claim that right-kind considerations and normative reasons are different kinds of things; indeed, as we noted earlier, one cannot make this claim without abandoning the "everybody's problem" response to the WKR problem. A more plausible view is that normative reasons are a type of right-kind reason. But we need a way to draw the line, other than by gesturing at a special kind of normativity that seems to smuggle in what they are trying to analyze.

## 2. THE FUNDAMENTALIST'S DILEMMA

So far our conclusions threaten everyone who wants to analyze normative notions in terms of normative reasons; given how plausible reasons-based analyses are of certain notions, our conclusions so far threaten just about everyone. Just as everyone has to differentiate between favorers of the right-kind from favorers of the wrong-kind, everyone has to differentiate favorers of the right-kind that are robustly normative from favorers of the wrong-kind that are only anemically normative. We promised, however, to raise deep worries for fundamentalism. Here we go.

We think the RKR problem raises a deep worry for fundamentalists. This is because fundamentalism imposes tight constraints on how the problem can

be solved, and these constraints make it quite unclear how fundamentalists can solve the RKR problem at all.

Recall what one must do in order to solve the RKR problem. First, one must tell us which activities and attitudes are normatively relevant. Which ones are such that right-kind reasons for them are robustly normative? Second, one must tell us what grounds the distinction between activities that are normatively relevant and activities that are not—i.e., what these activities have in common in virtue of which they are robustly normative. One should not rest content with a mere list.

These tasks are *particularly* hard for fundamentalists because it seems clear that the fundamentalist cannot appeal to *anything genuinely normative* to solve the problems. After all, there are only two options for a fundamentalist who appeals to normative properties to solve the problem: either she can appeal to something other than normative reasons or she can appeal to normative reasons. To appeal to some non-reason normative property would call fundamentalism into question. It would then be plausible that that property is at least as fundamental as reasons.

To appeal to normative reasons will not help, since what we want, precisely, is a story about which right-kind reasons are genuinely normative reasons. To appeal to normative reasons at this point would be to smuggle genuine normativity in through the back door, seemingly independently of reasons. But part of the job description of right-kind reasons is to analyze such normativity.

At this point things are looking pretty grim for the fundamentalist. It is unclear what resources are available to the fundamentalist to solve the Normativity and Rationale problems and thus to solve the RKR problem. Without normative stuff, what is left? The story would have to be told in naturalistic terms. But it is hard to see what that story could be.

We call this problem the *Fundamentalist's Dilemma*. The only options for solving the RKR problem are to appeal to normative properties or natural properties; the former route seems to be closed off by the constraints of fundamentalism, and it is simply hard to see what the latter route would involve. Even if the RKR problem does not strike one as pressing on its own, these problems lead straightforwardly to the Fundamentalist's Dilemma, which is pressing indeed.

### 3. POSSIBLE SOLUTIONS TO THE FUNDAMENTALIST'S DILEMMA

Is the Fundamentalist's Dilemma crippling? We hope not, since we are ourselves attracted to fundamentalism. In this section, we will consider some possible re-

sponses to the dilemma and indicate what we take to be the most promising solution.

Before we consider the options, it will be helpful to situate the options vis-à-vis *reduction*. Unsurprisingly, we have seen that fundamentalists cannot analyze normative RKR's in terms of something else that is normative. To do this would be to give up on fundamentalism. Hence, fundamentalism is incompatible with a certain *type* of reduction. This is a reduction of normative RKR's to something else that is *normative*. Despite this, fundamentalism is compatible with another type of reduction. This a reduction of normative RKR's to something *descriptive* or *natural*. So the fundamentalist has two options: either they think that normative RKR's can be reduced to something descriptive or natural, or they think that they are irreducible. We will call the latter view *reasons primitivism*.

As you might expect, there is more to say about reductive views than about primitivist views. Primitivist views, in effect, hold that normative RKR's are ground-floor fundamental. They are absolutely basic building blocks. There is not much one can do to positively develop this view. As even its proponents acknowledge, it is to be defended negatively—it is vindicated only by showing that all other views fail.<sup>19</sup> That said, below we will suggest that in fact something positive can be said by the primitivist about the Fundamentalist's Dilemma. Before we get to that, we will discuss two reductive accounts.

### 3.1. Humeanism

The Humean maintains that facts about normative RKR's reduce to facts about one's desires. Following Schroeder, we can say that, for the Humean, normative reasons for  $A$  to  $\phi$  are facts that explain why  $\phi$ -ing promotes at least one of  $A$ 's desires.<sup>20</sup> If we combine this with Activity, we can get an analysis of normative RKR's:

*Humean Activity:* What it is for  $r$  to be a normative RKR for  $A$  to  $\phi$  is for  $r$  to be a reason shared by everyone engaged in  $a$  and for  $r$  to explain why  $\phi$ -ing promotes some of  $A$ 's desires.

Humean Activity is a conjunctive account of normative RKR's. Interestingly, given a *prima facie* appealing assumption about being engaged in activities, the Humean can hold that in fact Activity is the correct analysis of normative RKR's.

The assumption is this: necessarily, if you are engaged in an activity  $a$ , then you have at least one desire that will be promoted if you do what everyone en-

19 See Parfit, *On What Matters*; Scanlon, *Being Realistic about Reasons*; and especially Korsgaard, *The Sources of Normativity*.

20 Schroeder, *Slaves of the Passions*.

gaged in *a* has a reason to do in virtue of being engaged in *a*. If this assumption is true and the Humean view is true, then it turns out that all RKR's are normative reasons. That is, it turns out we were wrong above when we insisted that some standards of correctness are normatively anemic. For the Humean, any standard of correctness can be normatively full-blooded for some agent *A* just in virtue of *A* engaging in an activity governed by that standard.

By making this move, the Humean fundamentalist does not solve the RKR problem, but rather dissolves it. This is because the problem arises only given the claim that not all RKR's are normative reasons. The Humean who makes good on the assumption thinks that all standards of correctness generate normative RKR's for those agents engaged in the relevant activities. So they have no burden explaining which activities generate normative RKR's, nor any burden explaining why the activities that generate normative RKR's do so. They presumably meet the latter burden just in virtue of vindicating Activity.

The ease with which the Humean view dissolves the problems is a serious virtue of the view. This virtue, as far as we know, has never been articulated (although we suspect Schroeder is well aware of it). However, this virtue should carry the day only if we accept both the Humean view of reasons and the assumption about engaging in activities (which we will call the *Engagement Assumption*). Unfortunately, both are seriously questionable. Furthermore, we think that the triad of views consisting of the Humean view, the Engagement Assumption, and Activity is implausible despite its ability to dissolve the problem. We will defend each of these claims in turn.

Let us start with the Engagement Assumption. While the Engagement Assumption does have initial plausibility, we think that it does not stand up to scrutiny. It seems initially plausible because many activities are essentially *intentional*. For those activities, it is plausible that someone engaged in that activity will necessarily have a desire promoted by conforming to the standards of the activity. That desire will be the desire associated with one's intention to engage in the activity.

The problem is that not all activities relevant to RKR's are essentially intentional.<sup>21</sup> Consider believing. One can engage in the activity of believing without intending to so engage. At some point in our development we come to have the capacity to believe, and at that point we just start believing. An intention to have

21 Schroeder concedes this point. He tries to get out of the problem by arguing that for each nonessentially intentional activity, there will be some essentially intentional activity that correlates with it (e.g., with admiration he suggests emulation). Sharadin shows that this strategy will not work ("Schroeder on the Wrong Kind of Reasons Problem for Attitudes" and "Reasons Wrong and Right").

beliefs is not necessary. Thus, it seems that believing is not an essentially intentional activity and so it is far from obvious that everyone has a desire that would be promoted by conforming to the constitutive standards of belief in virtue of engaging in it. So the Engagement Assumption does not look airtight. In fact, it looks like it undergeneralizes.

Furthermore, it is intuitively plausible that some standards are not genuinely normative. Of course, it is no secret that Humeanism has revisionary implications. And much has been said on both sides of the debate. We do not expect to settle the debate here. Nevertheless, the point can be put in a special way here. Consider certain ritualistic killings. Given that they are ritualistic, there are standards of correctness governing them. It is possible to perform the killings in a way that deviates from these standards. So a WKR problem can arise for these ritualistic killings. Are the RKRS relative to those standards normative RKRS for those (intentionally) engaged in those activities? Plausibly not. This would make engagement in horrid activities bizarrely self-vindicating.

We know that there are possible ways out for the Humean. In fact, we recognize that we have added ammunition to their arsenal by showing that they have a powerful reply to the Fundamentalists' Dilemma. Nonetheless, the triad of views is not all that plausible. This is for two reasons. First, it does not look like the Engagement Assumption that is needed to get the reply to work is going to get the right predictions out of Activity. That is, it does not look like the kinds of activities relevant to Activity are essentially intentional activities. But that is what is needed in order for the Humean view to show that all RKRS are normative reasons, which is what the Humean needs in order to dissolve the Fundamentalists' Dilemma. The second problem is a deeper problem with Humeanism: it generates too many reasons and generates them too easily.

### 3.2. *Constitutivism*

A second reductive response is a version of constitutivism. It attempts to solve the Normativity and Rationale problems in one fell swoop by telling us that the activities for which right-kind reasons are normative reasons are the activities any agent must be engaged in *qua* agent. It defuses the Fundamentalists' Dilemma by noting that the line it draws to distinguish between right-kind reasons that are normative and right-kind reasons that are not is a descriptive line. In this way, constitutivism provides an elegant naturalist reductive solution to the fundamentalists' problems.

Let us walk through the solution at a slower pace. Let us start with the Normativity Problem. The constitutivist might start by pointing out that not all activities are created equal. In particular, some activities seem essential to our

agency. We can call these activities the activities essential to *deliberation*. When we deliberate practically, we deliberate about what to want, what to intend, and what to do. When we deliberate theoretically, we deliberate about what to believe and how confident we should be about various claims. To engage in deliberation is *just what it is* to be an agent. This is a natural or descriptive claim about what agency is.

The constitutivist holds that we can reduce facts about normative RKR to facts about agency. In particular, we can analyze normative RKR in terms of the activities that are constitutive of agency:

*Constitutive Activity*: What it is for  $r$  to be a normative RKR to  $\phi$  relative to some activity  $a$  is for  $r$  to be a reason shared by everyone engaged in  $a$  in virtue of being engaged in  $a$  and for  $a$  to be an activity constitutive of agency.

This is a constitutivist view twice over because it understands RKR in terms of the constitutive norms of activities and it understands *normative RKR* in terms of what is constitutive of agency. We will call the former view *constitutivism about RKR* and the latter position *constitutivism about normative RKR*.<sup>22</sup>

This is an interesting analysis. It is also predictive once it is supplemented with views about the activities constitutive of agency. And there are already prominent views in the literature about particular activities, with believing being the most well-developed.<sup>23</sup> The plausibility of these views provides indirect evidence for Constitutive Activity. Consider belief again. Believing is plausibly on the list of normatively relevant activities. It is plausible that right-kind reasons for belief are normative reasons. The standard of correctness for belief is, one might think, surely not on all fours with norms of etiquette or some specific set of instructions for cooking *cacio e pepe*. But what could distinguish the standard of correctness for belief from these other standards other than the fact that “playing the belief game” is a constitutive feature of agency? If the constitutivist story provides a credible answer—perhaps the only obvious answer—to this question about the normativity of belief’s standard of correctness, explanatory parsimony recommends telling the same story elsewhere.

This version of constitutivism has an obvious solution to offer to the Rationale Problem. What unifies the activities on the list is the fact that they are all

22 We were implicitly relying on constitutivism about RKR when we laid out Activity above. We are making it explicit here mostly because it will be helpful in the next subsection.

23 See, e.g., Velleman, “On the Aim of Belief”; Shah, “How Truth Governs Belief”; Sosa, *Judgment and Agency*; Wedgwood, *The Value of Rationality*. For discussion of intention, see Shah, “How Action Governs Intention.”

constitutive of agency. One might find appeals to agency normatively irrelevant, but all should agree that this type of appeal to what is constitutive of agency is interestingly unified.

There are, of course, challenges for constitutivists. Our goal is not to provide a comprehensive defense of constitutivism here, but we will consider two challenges. The first is the most notable general challenge to constitutivism and is often thought to plague all versions of the view. This is the “schmagency” challenge.<sup>24</sup> This challenge, though, does not obviously apply to the type of constitutivist view appealed to here. This is because the force of the challenge is felt only when certain background assumptions are made about *authority*. These background assumptions are made by some constitutivists. The most notable example is Christine Korsgaard.<sup>25</sup> Korsgaard is interested in employing constitutivism in a particular context. The operative question in this context is a question about the normative authority of morality—why be moral? Given Korsgaard’s unorthodox views about the function of moral concepts, she has a very demanding view about what it takes to vindicate the normative authority of morality.<sup>26</sup> Any satisfactory answer, for Korsgaard, will strike any agent inquiring about what to do as a satisfactory answer. This standard demands from a successful answer that it convinces a skeptic—a tall order indeed.

These background views open Korsgaard to a particularly damning version of the schmagency problem. For it seems entirely intelligible to skeptically challenge the normative relevance of what is constitutive of agency. The normative relevance of agency is an open question like (nearly) anything else. One can intelligibly ask why one should not opt out of agency in order to become a schmagent. What this shows is that Korsgaard’s demands for a successful answer cannot be met. She has a schmagency problem.

But not all constitutivist views have the ambitions of Korsgaard’s view. We are assuming that the proponent of Constitutive Activity is not in the business of convincing the skeptic *via* Constitutive Activity. Moreover, we recommend to the proponent of Constitutive Activity that she not go in for answering whether there is some special reason for being an agent rather than a schmagent. That question has a false presupposition—*viz.* that constitutivism is committed to thinking that there is such a reason.

Sometimes we have no reasons to be agents any longer. Sometimes we should kill ourselves. Other times we might not have particularly compelling reasons to

24 This challenge was originally articulated in Enoch, “Agency, Shmagency.” See also Enoch, “Shmagency Revisited.”

25 See especially Korsgaard, *The Sources of Normativity*.

26 See Korsgaard, *The Sources of Normativity*, ch. 1, for the (complicated) details.

be an agent, but, nevertheless, we are agents. We thus have all sorts of normative reasons, per Constitutive Activity. The fact that, in these cases, agents do not have particular strong reasons to be agents is neither here nor there. Without appealing to claims about authority, it is hard to generate pressure to think otherwise. The constitutivist should deny those overly ambitious claims about authority.<sup>27</sup>

This challenge also strikes us as weak in the context of the Fundamentalist's Dilemma. Reasons fundamentalism had a lot going for it. It offers a beautiful account of the internal structure of the normative. Constitutivism provides the obvious resolution of the Fundamentalist's Dilemma. It is plausible on first-order grounds that the list of activities suggested by constitutivism is the list of activities for which right-kind reasons are normative reasons. But if one is already attracted to reasons fundamentalism, there seems to be no deep explanation of why that list is the right list other than the explanation given by the constitutivist. If reasons fundamentalism is true, the Fundamentalist's Dilemma must have an answer. Perhaps the best argument for constitutivism is that it provides this answer.

The second challenge we will consider is a more local worry for Constitutive Activity. The core of the worry is that Constitutive Activity undergenerates because not enough activities are constitutive of agency to generate enough normative RKR's. We can see the worry by thinking about particular attitudes. Admiration seems to be the sort of attitude we have normative RKR's to have. According to Constitutive Activity, this is true only if the activity that governs admiration is an activity that is constitutive of agency. But this is far from clear. It seems plausible that we would be the same sort of creature even if we did not have the capacity for admiration.

We think that this is a serious challenge to Constitutive Activity. Further, the point is obviously not limited to admiration. There are likely many other reactions that intuitively are reactions we have normative RKR's to have even though it is not obvious that the activities associated with those reactions are constitutive to agency. We will call these *Normative but Optional Activities*. We cannot hope to meet these challenges here, but we will sketch out two strategies for meeting them.

The first strategy is simple. It insists that some particular Normative but Optional Activity is in fact constitutive of the sort of agents we are. We will call this the *Doubling Down Strategy*. In order to apply the Doubling Down Strategy in all cases, it seems to us that one would need to adopt a fragile conception of what constitutes the sort of agency normal agents exemplify. On such a fragile understanding, it is easy to cease being the sorts of agents normal agents are. In fact, it might be that some actual human beings fail to be the kind of agents that

27 For a similar response, see Smith, "The Magic of Constitutivism."

normal agents are (e.g., perhaps psychopaths fail to be the sort of agents normal adults are). Such a fragile view will have some implications that some will find unappealing, but they do not strike one as wildly implausible on their face.<sup>28</sup>

Furthermore, one need not accept the fragile view of agency in order to use the Doubling Down Strategy in a more local way. For example, we are actually skeptical that admiration is a Normative but Optional Activity even if one accepts a robust account of the constitution of our agency. This is because we find it plausible that admiration is a central reaction to our moral agency, which is a constitutive part of our agency even if you have a robust view.<sup>29</sup>

The second strategy insists that the purported Normative but Optional Activities are reducible to reactions that are obviously constitutive of our agency, with the obvious contenders being belief and desire. We call this the *Reduction Strategy*. In order to apply the Reduction Strategy in full generality, one would likely commit oneself to a sparse philosophical psychology. We note, though, that the exact sort of project that seems needed is already rigorously defended by several prominent philosophical psychologists.<sup>30</sup>

Of course, just like with the Doubling Down Strategy, one might apply the Reductive Strategy more locally. When it comes to admiration, for example, one might argue that admiration reduces to a bundle of beliefs and desires. To take an oversimplistic view, one might hold that admiration is the bundle of a belief that the object of admiration is desirable and the desire to emulate the admired.<sup>31</sup> If such a reduction were true (or a suitably more complex version), then it would be plausible that RKR for admiration are normative RKR. For, if the reduction were true, then RKR for admiration would just be RKR for having certain beliefs and desires. Believing and desiring are clearly constitutive of agency and thus RKR for believing and desiring are normative, per Constitutive Activity.<sup>32</sup>

28 Some evidence for this is that Korsgaard adopts a very fragile view of one's practical identity in recent work (see especially *Self-Constitution*).

29 See Zagzebski, "Admiration and the Admirable"; Irwin, "Nil Admirari?"

30 E.g., Arpaly and Schroeder, *In Praise of Desire*; and Sinhababu, *Humean Nature*.

31 For discussion, see Zagzebski, "Admiration and the Admirable."

32 Interestingly, this does not quite show that *admiration* is constitutive of agency. For even if the reduction is true, it seems possible to be the type of agents we are without having admiration—i.e., without having the bundles of beliefs and desires that constitute admiration. Even if it does not show that admiration is constitutive of agency, it does show that the RKR for admiration are normative RKR for those who do admire. This might seem to threaten the letter of Constitutive Agency. We do not think that it does. The right way to think of it is that if the reduction is true, then the activities that are relevant are believing and desiring. So if admiration is not constitutive but the reduction is true, admiring is just not the activity that we plug into the analysis in order to determine what the normative RKR are for

These are the strategies we think defenders of Constitutive Activity should use to combat cases of purported Normative but Optional Activities. While we are skeptical that either can be used in full generality, we are hopeful that local uses of each strategy can stave off the most pressing worries. What we have hoped to show here is that with more work these objections can be overcome and thus that we should continue to think of Constitutive Activity as a live option.

While this argument on its own might not sell constitutivism, in combination with reasons fundamentalism and the need for an answer to the Fundamentalists' Dilemma strike us as providing a serious argument for constitutivism. At the very least, the disjunction *either constitutivism is true or reasons fundamentalism is false* merits high credence.

### 3.4. A Louder Quietism

We will end by discussing a non-reductive view. We end with this view not because we think it is the most plausible—far from it—but rather because our discussion of the reductive views will help us see how a quietist position can be a bit more informative than some quietist views.<sup>33</sup> Just to be clear, however: this quietist view is a quietist view. That is, it holds that there is nothing informative that can be said at the analytic level about why the normative RKR's are normative RKR's. They cannot be analyzed in terms of something normative—hence fundamentalism. They cannot be analyzed in terms of something natural or descriptive—hence primitivism. This much is just run-of-the-mill quietism.

Nevertheless, according to the louder quietism we have in mind, the RKR problem does have a solution of a certain kind. There is a compact list to be given of the attitudes for which right-kind reasons are normative reasons, and there is a principled story to be told about what the members of this list have in common. But according to our quietism, the principled story to be given is at bottom a piece of *first-order normative theorizing*, one that is supported by systematizing our intuitions about what is a reason for what. It does not tell us anything about how the normative is grounded in the natural or descriptive, *pace* the constitutivist and Humean.

---

admiring. We need to plug in believing and desiring. Thanks to an anonymous referee for prompting us to say more here.

33 We have in mind Parfit and Scanlon in particular. We recognize that some quietists might be immune to everything we say below. We also recognize that Parfit and Scanlon have engaged in the sort of first-order project we discuss when it comes to parts of the normative domain, with morality being the domain they have focused on the most. The view we sketch here is supposed to play a similar role but at a much more general level.

To make this view more concrete, we will sketch one way that it could be developed, though it is only one that we mention for the sake of illustration. This version of the view gives the same list of normatively relevant activities that the constitutivist view gives in solving the Normativity Problem, and even agrees with the constitutivist about what the elements of that list have in common. But it provides a first-order rationale for the latter claim. The most natural way for quietists to do all of these things is by becoming constitutivists about RKR. That is, by thinking that the reasons that are shared by everyone engaged in an activity in virtue of being engaged in that activity are the reasons that have some connection to the constitutive standards of that activity. Normative reasons are then taken to be RKR generated for the activities constitutive of agency, where this is made as a first-order claim. It is a first-order view with some plausibility, as the literature on how to derive a truth-norm on belief by appeal to constitutive norms illustrates.<sup>34</sup>

To see the main point more clearly, think about the Moorean view about goodness and pleasure. The Moorean view about goodness is a primitivist and nonnaturalist one. It holds that goodness is an unanalyzable property that is different in kind from the natural properties. Nevertheless, the Moorean thinks that the only states of the world that had this property are pleasure states. This does not mean that the property of being good is the same as the property of being pleasant. Rather, it is just to say that *the* good is pleasure. So a certain kind of first-order hedonism is, of course, compatible with primitivist nonnaturalism.

The nonnaturalist constitutivist view we have imagined is like the Moorean version of hedonism. It holds that the right-kind reasons (the reason analogue of the good) are the facts that are tied to constitutive standards in a certain way. This does not mean that the property of being a normative reason is the same as the property of being a fact that stands in the right relation to constitutive standards. All that is posited is a necessary correlation.<sup>35</sup>

This version of quietism is more interesting than run-of-the-mill quietism be-

34 Again, see, e.g., Shah, "How Truth Governs Belief"; and Velleman, *The Possibility of Practical Reason*.

35 The existence of Moorean hedonism does not show that there are not intelligible metaphysical questions to ask about what the good *is*. Of course, Mooreans deny that these questions have substantive answers. But it would be a bad move to claim that there are no intelligible metaphysical questions to ask about the nature of goodness simply because one can hold that the independent property of goodness necessarily correlates with pleasure. Analogously, it would be a bad move to hold that we cannot intelligibly ask the metaphysical question of what RKR are normative RKR just because one could hold that the independent property of being a normative RKR necessarily correlates with the property of being a consideration that is a reason shared by all engaged in an activity constitutive of agency.

cause it gives a unifying view about which facts are normative RKRS. It does not give us a metaphysical explanation of *why* they are like this. Nor does it tell us that the *kind of things* normative RKRS are is the same as the kind of things tables, chairs, and electrons are. Nevertheless, it tells a unifying story about which facts are the normative RKRS.

While this view has a bit more predictive power than run-of-the-mill quietism, it is still anchored in a strange bit of metaphysics. It posits a *brute* necessary correlation between RKRS and constitutive standards.<sup>36</sup> Furthermore, this is an essential feature of the view. This is what makes it a quietist view. This makes it stark that the “solution” it offers to the problems is not much of a solution. The best the quietist can do is change the subject from the second-order question about the metaphysics of normative RKRS to the first-order question about what reasons there are. They remain silent about the second-order question except to say that it is not a question worth trying to answer.

### 3.5. *Sizing Up the Options*

We have explored three replies to the Fundamentalists’ Dilemma. The first two were reductive. They attempt to analyze normative RKRS in terms of something natural. This is compatible with fundamentalism because fundamentalism is merely the claim that normative RKRS are *normatively* fundamental. The last view was a quietist nonnaturalist view. It held that normative RKRS are not analyzed at all. They are basic building blocks. Nevertheless, we saw that the quietist could borrow some machinery from the reductive constitutivist account in order to have a bit more predictive power.

We think the best option is the reductive constitutivist one. Constitutivism about RKRS is, we think, the most natural way of developing Activity. Moreover, it is striking how plausible it is that the activities that are constitutive of agency are the activities that would naturally go on one’s list of activities that are nor-

36 One might wonder if this only looks odd because we have been assuming that the solution to the Normativity Problem maintains that all the RKRS for the robustly normative activities are normative RKRS. If one were to deny this, would there be a brute necessary connection? Well, it is right that the brute connection we are complaining about here is a brute connection between normative RKRS and the considerations that provide evidence that the standards of correctness of certain activities constitutive of agency are met. But even quietists who deny that all the RKRS of constitutive activities provide normative RKRS posit a brute connection. Any quietist who holds that there is a correlation between some descriptive facts and some normative facts posits a brute connection. So you cannot get out of a brute connection by giving up our assumption. By giving it up, you will certainly have a less elegant theory, though. Pending a surprising explanation of why some RKRS associated with constitutive activities are not normative, this view seems to lose the elegance of the view we consider in the text. Thanks to an anonymous referee for raising this point.

mative in a full-blooded sense. This is obviously far from decisive, but we are intrigued by it.

Furthermore, the constitutivist view has serious virtues compared to the other two views. It has a principled way of avoiding the explosion of normative RKR's that the Humean is saddled with. The constitutivist view lets us hold on to the plausible thought that not all standards are created normatively equal. Some of them are normatively irrelevant for everyone. The Humean, by contrast, holds that any standard can become normatively relevant for an agent just in virtue of that agent engaging in an activity governed by those standards. That makes full-blooded normativity too cheap.

The constitutivist view has a serious explanatory advantage over the quietist view. The quietist view posits a brute metaphysical necessity between normative RKR's and the constitutive standards of the activities that are constitutive of agency. Brute metaphysical necessities are born of desperation. We should avoid desperation if we can. The reductive constitutivist view allows us to avoid desperation because it offers an explanation of the metaphysical necessity. There is a necessary correlation between normative RKR's and the constitutive standards of the activities constitutive of agency because *what it is* to be a normative RKR is to be an RKR relative to an activity that is constitutive of agency.

These are serious virtues of the constitutivist view. They are obviously not fully decisive. We have not surveyed all the possible solutions to the Fundamentalists' Dilemma. Nor have we seriously considered the option of abandoning fundamentalism because of the dilemma. Our interest has been in clearly articulating the dilemma and seeing whether there are good routes out of the problem. The constitutivist route is the best option we have thought of. Whether this a failure of our imaginations is yet to be determined.

#### 4. CONCLUDING MORALS

Let us take stock of the key morals. The first key moral is that the generality of the WKR/RKR distinction shows that there is another problem that is structurally similar to the WKR Problem. This is the RKR Problem. The generality of the WKR/RKR distinction gives rise to the RKR problem because the generality of the problem shows that not all RKR's are *normative* reasons. Accordingly, those of us who analyze complex normative notions in terms of normative reasons—i.e., everyone—cannot pick out our analysans just by providing a principled distinction between WKRs and RKR's. We also need a story about which activities/attitudes are the ones for which RKR's are normative reasons. But this task is highly nontrivial. Further, this task is especially difficult for the fundamentalist. We

demonstrated this via the Fundamentalist's Dilemma. Fundamentalism blocks one from appealing to anything normative to solve the RKR problem. This forces the fundamentalist to pick between quietism and reductionism.

The moral of the story is that, to properly solve the RKR problem, the fundamentalist is forced into taking a stand about the metaphysical nature of normative reasons. That is, she has to take a stand about whether naturalism or non-naturalism is true. In some ways this is unsurprising once you go for the thought that *something* needs to be said about what the basic building blocks are like. And the WKR problem clearly forces the fundamentalist to say something about this. Relying on the "counts in favor" idiom does not cut it. But it is obvious that the fundamentalist cannot appeal to anything normative to explain which facts that count in favor are her building blocks. So if she is going to say anything with explanatory power, she will need to appeal to something natural.

The surprising twist is that she will have to do more than solve the WKR problem. She will also have to solve the RKR problem. As we saw, this problem breaks down into the problem of telling us which activities/attitudes are the ones for which RKR are normative reasons, and the problem of telling us why. It is here, we suggested, where the fundamentalist will take a stand over naturalism. We have suggested that the stand she should take is in favor of a form of reductive constitutivism. This view offers the best balance of explanatory power and extensional adequacy. And those are the right kind of reasons for accepting a theory.<sup>37</sup>

*University of Pennsylvania*  
 erlord@phil.upenn.edu

*University of Southampton*  
 k.l.sylvan@soton.ac.uk

#### REFERENCES

Arpaly, Nomy, and Timothy Schroeder. *In Praise of Desire*. Oxford: Oxford University Press, 2014.

Baker, Derek Clayton. "Skepticism about Ought Simpliciter." In *Oxford Studies*

37 For helpful feedback, we would like to thank two anonymous referees for this journal and audiences at the Southern Normativity Group, SLACRR, and the Uppsala Language and Metaphysics of Normativity conference, which included Matti Eklund, Daniel Fogal, Amelia Hicks, Simon Kirchin, Stephanie Leary, Kathryn Lindeman, Conor McHugh, Pekka Väyrynen, Jonathan Way, Daniel Whiting, and Jack Woods. We would also like to acknowledge our debt to Nate Sharadin and Barry Maguire for crucial inspiration at the beginning.

- in *Metaethics*, vol. 13, edited by Russ Shafer-Landau. Oxford: Oxford University Press, 2018.
- Copp, David. "Moral Naturalism and Three Grades of Normativity." In *Morality in a Natural World: Selected Essays in Metaethics*, 249–83. Cambridge: Cambridge Studies in Philosophy, 2004.
- Dancy, Jonathan. *Ethics without Principles*. Oxford: Oxford University Press, 2004.
- D'Arms, Justin, and Daniel Jacobson. "The Moralistic Fallacy: On the 'Appropriateness' of Emotions." *Philosophy and Phenomenological Research* 61, no. 1 (July 2000): 65–90.
- . "Sentiment and Value." *Ethics* 110, no. 4 (July 2000): 722–48.
- Enoch, David. "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Agency." *Philosophical Review* 115, no. 2 (April 2006): 169–98.
- . "Shmagency Revisited." In *New Waves in Metaethics*, edited by Michael S. Brady, 208–33. Basingstoke, UK: Palgrave-Macmillan, 2011.
- Foot, Philippa. "Morality as a System of Hypothetical Imperatives." *Philosophical Review* 81, no. 3 (July 1972): 305–16.
- Irwin, T. H. "Nil Admirari? Uses and Abuses of Admiration." *Aristotelian Society Supplementary Volume* 89, no. 1 (June 2015): 223–48.
- Korsgaard, Christine M. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press, 2009.
- . *The Sources of Normativity*. Cambridge, MA: Harvard University Press, 1996.
- McPherson, Tristram. "Against Quietist Normative Realism." *Philosophical Studies* 154, no. 2 (June 2011): 223–40.
- Parfit, Derek. *On What Matters*. Oxford: Oxford University Press, 2011.
- Plunkett, David, and Scott J. Shapiro. "Law, Morality, and Everything Else: General Jurisprudence as a Branch of Metanormative Inquiry." *Ethics* 128, no. 1 (October 2017): 37–68.
- Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen. "The Strike of the Demon: On Fitting Pro-attitudes and Value." *Ethics* 114, no. 3 (April 2004): 391–423.
- Scanlon, T. M. *Being Realistic about Reasons*. Oxford: Oxford University Press, 2014.
- Schroeder, Mark. *Slaves of the Passions*. Oxford: Oxford University Press, 2007.
- . "Value and the Right Kind of Reason." In *Oxford Studies in Metaethics*, vol. 5, edited by Russ Shafer-Landau, 25–55. Oxford: Oxford University Press, 2010.

- Shah, Nishi. "How Action Governs Intention." *Philosophers' Imprint* 8, no. 5 (July 2008): 1–19.
- . "How Truth Governs Belief." *Philosophical Review* 112, no. 4 (October 2003): 447–82.
- Sharadin, Nathaniel. "Reasons Wrong and Right." *Pacific Philosophical Quarterly* 97, no. 3: (September 2016): 371–99.
- . "Schroeder on the Wrong Kind of Reasons Problem for Attitudes." *Journal of Ethics and Social Philosophy* 7, no. 3 (August 2013): 1–8.
- Sinhababu, Neil. *Humean Nature: How Desire Explains Action, Thought, and Feeling*. Oxford: Oxford University Press, 2017.
- Skorupski, John. *The Domain of Reasons*. Oxford: Oxford University Press, 2011.
- Smith, Michael. "The Magic of Constitutivism." *American Philosophical Quarterly* 52, no. 2 (2015): 187–200.
- Sosa, Ernest. *Judgment and Agency*. Oxford: Oxford University Press, 2015.
- Velleman, J. David. "On the Aim of Belief." In Velleman, *The Possibility of Practical Reason*, 244–81.
- . *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000.
- Wedgwood, Ralph. *The Value of Rationality*. Oxford: Oxford University Press, 2017.
- Wodak, Daniel. "Mere Formalities: Fictional Normativity and Normative Authority." *Canadian Journal of Philosophy* (forthcoming).
- Woods, Jack. "The Authority of Formality." *Oxford Studies in Metaethics*, vol. 13, edited by Russ Shafer-Landau. Oxford: Oxford University Press, 2018.
- Zagzebski, Linda. "Admiration and the Admirable." *Aristotelian Society Supplementary Volume* 89, no. 1 (June 2015): 205–21.

## HELPING BUCHANAN ON HELPING THE REBELS

*Daniel Weltman*

MASSIMO RENZO has recently argued in this journal that Allen Buchanan's account of the ethics of intervention is too permissive.<sup>1</sup> Renzo claims that a proper understanding of political self-determination shows that it is often impermissible to intervene in order to establish a regime that leads to more self-determination for a group of people if that group was or would be opposed to the intervention. Renzo's argument rests on an analogy between individual self-determination and group self-determination. However, the analogy also points to crucial differences between the two kinds of self-determination. To make his argument work, Renzo must come up with a theory of self-determination that accounts for these differences without vitiating his argument, and it is not clear that this can be accomplished. In response to the differences we may in fact be pushed to adopt an account of self-determination that is more permissive with respect to intervention than even Buchanan's theory.

### 1. RENZO'S ARGUMENT AND THE ANALOGY BETWEEN INDIVIDUALS AND GROUPS

Buchanan argues that, for the sake of self-determination, it can make sense to intervene in the affairs of a state, and thus the right to self-determination is not as strong a barrier against intervention as one might have thought.<sup>2</sup> He addresses cases in which an oppressive regime is going to crush a democratic revolution or a revolution is going to create an authoritarian government. He argues that, because these are cases in which the state is not or will not be self-determining, it is not a violation of the state's right to self-determination if third parties intervene in order to ensure the creation of a democratic state that would be self-deter-

- 1 Renzo, "Helping the Rebels"; Buchanan, "The Ethics of Revolution and Its Implications for the Ethics of Intervention"; Buchanan, "Self-Determination, Revolution, and Intervention."
- 2 His targets include Michael Walzer (*Just and Unjust Wars*, 84–104), who claims that the only justifiable intervention outside of extreme cases like genocide is one that balances out the intervention of some other party, because anything else would violate self-determination.

mining.<sup>3</sup> More controversially, he claims that intervention could be justified in order to overrule a newly elected democratic government if this government is going to eliminate the opportunity for self-determination in the future. Because the right to self-determination is not a right to undercut the ability of others to self-determine, it does not violate a group's self-determination to prevent it from stopping another group (future inhabitants of the country) from exercising self-determination.<sup>4</sup> Or, it does violate its right, but this is a permissible violation for the sake of others.<sup>5</sup>

Renzo's main response is to point out that Buchanan assumes that we have to look at a group's present exercise of self-determination to decide whether its right to self-determination would be violated by intervention. If we widen our scope to look at the past and at what the group would want right now if it were self-determining, this can give us reason to think that intervention would impermissibly violate the group's self-determination, if in the past it expressed a wish not to be intervened in or if we can reasonably attribute such a desire to the group on the basis of its present values and preferences.<sup>6</sup> If either of these possibilities suggests that the group did not or does not consent to intervention, then it is an impermissible violation of that group's right to self-determination to intervene.

Renzo aims to make this plausible by giving the example of an individual. Say that you are a pacifist and Renzo takes control of you through hypnosis. If I intervene by killing Renzo and freeing you, even though in the past you told me you would not like me to do this or even though I can infer that, as a pacifist, you would not approve of this, then I have violated rather than protected your self-determination. This is true even though, at the moment, your self-determination is eclipsed because Renzo hypnotized you.

As a story about individuals, let us grant that this works fine. With respect to intervention in groups, though, it is not as clearly compelling, for two reasons: one practical and another theoretical.

First, the practical reason. As Renzo notes, Buchanan is particularly aware of the actual mechanisms by which revolutions and other opportunities for inter-

3 Buchanan, "The Ethics of Revolution and Its Implications for the Ethics of Intervention," 460–61.

4 Buchanan, "The Ethics of Revolution and Its Implications for the Ethics of Intervention," 462.

5 Buchanan, "The Ethics of Revolution and Its Implications for the Ethics of Intervention," 464.

6 Renzo, "Helping the Rebels," 231–32.

vention occur.<sup>7</sup> Buchanan's concerns are first and foremost about whether the right to political self-determination is an obstacle to intervention in actual cases. In these actual cases, typically it will be very hard to draw conclusions about what the group wanted in the past and what the group would want now. Take for instance Renzo's own example of Cyprus, which signed a treaty in 1960 authorizing other countries to intervene if necessary to reestablish the status quo.<sup>8</sup> Must we respect this treaty to respect Cyprus's present self-determination? Do the signatories to the treaty (the president and vice president of Cyprus at the time) represent present-day Cyprus (or even past Cyprus)? Matters become even more complicated if we are not dealing with explicit treaties signed by democratically elected leaders. In many of the cases of intervention that we are concerned with, there will be no previous explicit statements about intervention. Even if there are, they will have been made by leaders who were not elected and thus may not have plausibly represented anyone in the past, let alone in the present. Similar issues apply to inferring a group's present desires. How confident can we be that a group actually would or would not desire intervention? The American government famously misjudged how sanguine Iraqis would be at the prospects of an American invasion. Although we might think that this particular judgment was obviously flawed, the general point is that there will often be significant difficulties attached to inferring a group's present preferences absent mechanisms like a fair democratic vote.

Beyond these practical problems lie theoretical difficulties. Groups are different from individuals in many ways. As Renzo alludes to with his reference to Derek Parfit and psychological connectedness, if there is any doubt about the link between your decision in the past and your status in the present, we would not want to give weight to that decision in the past.<sup>9</sup> For people, there is rarely this kind of doubt, but for groups the doubt is omnipresent.<sup>10</sup> As Thomas Jefferson pointed out, groups change over time, and this change causes us to wonder about whether a group's past decisions still bind the present:

It is now forty years since the constitution of Virginia was formed....  
Within that period, two-thirds of the adults then living are now dead.  
Have then the remaining third, even if they had the wish, the right to hold

7 Renzo, "Helping the Rebels," 223.

8 Renzo, "Helping the Rebels," 231.

9 Renzo, "Helping the Rebels," 230; Parfit, *Reasons and Persons*, pt. 3.

10 Of course, if this doubt is omnipresent for people, then Renzo's case is even more tenuous. So, for the sake of the argument, we will assume that the individual case is not problematic. I thank an anonymous reviewer for suggesting I say more about this.

in obedience to their will, and to laws heretofore made by them, the other two-thirds, who, with themselves, compose the present mass of adults?<sup>11</sup>

This concern for the changing makeup of groups is present in Buchanan's argument, which is focused to a large degree on the importance of using intervention to prevent present people from precluding the self-determination of future people.<sup>12</sup> Similar concerns apply to using the decisions of a group in the past to determine whether there is a right to intervene in the group now.

The differences between groups and individuals are not limited to the fact that groups change composition over time. Groups are also less clearly delineated than individuals. Individuals, at least in normal cases, have one human body and one mind. Not so for groups. Let us assume that I can easily ascertain which people an individual comprises (there will be just one person).<sup>13</sup> Figuring out who a group comprises is much less straightforward. Renzo takes it for granted that we know which groups we are talking about and thus whose self-determination is at stake. But we do not.<sup>14</sup> The Cyprus example is again illustrative. Cyprus has long been divided between Greek Cypriots and Turkish Cypriots. This has occasioned violence and acrimony (including an invasion by Turkey, a partition of the island of Cyprus, and continued Turkish occupation of northern Cyprus), and some very different opinions on what ought to happen in Cyprus. For the purposes of figuring out whose right of self-determination to respect, how many groups do we have? Is there one group of Cypriots, which, according to the principle of majority rule, wants Turkey gone? Are there two groups, one of which wants Turkey to leave and the other that does not mind Turkey's presence? Obviously how we draw the borders for group membership matters quite a bit in terms of figuring out what respect for self-determination entails. Would it respect self-determination if Turkey left, or if Turkey remained?

For these reasons, then, things are not as simple as Renzo presents them. It is straightforward to figure out what I wanted in the past and what I would want now if I could get my way. It is nowhere near as straightforward to do this for groups. So, unless we say more, Buchanan's original arguments withstand Renzo's challenge: for the sake of a group's self-determination, it can make sense to intervene without consent from a group, because the alternatives Renzo raises are not apposite. Unlike the individual case, we cannot straightforwardly

11 Jefferson to Samuel Kercheval, July 12, 1816, 7–8.

12 Buchanan, "The Ethics of Revolution and Its Implications for the Ethics of Intervention," 462, 464.

13 Again, this may be false, but if it is, that is worse for Renzo.

14 For a more complete defense of this point, see Weltman, "Who Is the Self in Self-Determination?"

talk about what some group wanted in the past or would want right now in order to figure out whether there is a right to intervene.

2. WIDENING THE SCOPE OF PERMISSIBLE  
INTERVENTION: A UTILITARIANISM OF RIGHTS

This leaves us with a worry, though. We have just seen how, depending on how we delineate groups, we get very different results in terms of the right to self-determination. If there is one group in Cyprus, Cyprus wants Turkey to leave. If there are two, then one wants Turkey to stay. So, we first need an account of what the relevant groups are before we start to worry about their right to self-determination. Before we can figure out whether intervention violates a group's self-determination, we have to figure out what group or groups we are dealing with. This worry applies to any argument that relies on any sort of right to group self-determination, so it threatens Buchanan as much as it does Renzo.

One solution would be to say that this may be an issue in principle but not in practice. Practically speaking, we know when we are dealing with one group rather than two or more. Unfortunately this seems false in many of the cases in which intervention is a pressing question. Deeply divided societies facing conflicts like revolution are precisely the ones where intervention will be a live topic. Turkey's intervention in Cyprus, for instance, cannot be understood without seeing how Cyprus could coherently have been viewed as a society consisting of at least two groups, rather than one unified society.

Another solution would be to adopt a more specific theory of self-determination than the ones adverted to by Buchanan and Renzo. One option is to say that we only care about self-determination for nations, which are groups that share an encompassing culture.<sup>15</sup> (This is clearly not what Buchanan has in mind. He argues that his account floats free of this issue and is compatible with whatever account of the group we pick.<sup>16</sup>) I cannot adjudicate the entire nationalist versus non-nationalist debate here. What I can say is that if the nationalist account is true and nations are the only sorts of groups with a right to self-determination, it is still not obvious that the past decisions of a nation's self-proclaimed leaders can bind its future members, or that we can easily tell what a nation thinks about some topic if it is not presently able to rule itself. We should worry that if people are not able to vote or otherwise freely express themselves, privileged elites in

15 See Margalit and Raz, "National Self-Determination," 442–47.

16 Buchanan, "Self-Determination, Revolution, and Intervention," 470. Elsewhere, Buchanan rejects nationalistic theories of self-determination. Buchanan, *Justice, Legitimacy, and Self-Determination*, 344.

the nation will claim to speak for the group when in reality they speak for themselves, with the result that elites may reject interventions that the nation more generally would welcome. So we should keep looking.

There are many other options. There is self-determination as:

- a consequence of freedom of association;<sup>17</sup>
- a consequence of the interests that individuals have in being ruled by institutions that reflect their priorities;<sup>18</sup>
- based on the value of individual autonomy;<sup>19</sup>
- based on individual interests in establishing and revising laws and public norms;<sup>20</sup>
- based on the value of collective autonomy;<sup>21</sup>
- ownership of state institutions due to past contributions.<sup>22</sup>

Renzo, like Buchanan, does not commit to any particular theory.<sup>23</sup> Each of these theories has a different way of picking out the relevant groups and thus telling us whether, say, Cyprus is one “self” or two “selves” when it comes to self-determination, at least in principle. At the very least, Renzo needs to provide an account that is convincing in light of the practical concerns Buchanan has in mind, because in messy cases like Cyprus we cannot take for granted that we know how many “selves” we are dealing with or how stable over time those “selves” are. There are, moreover, reasons to worry that any account of the “self” can provide a good theoretical answer to questions of potentially divided selves, or a good practical answer in the messy cases of potential intervention.<sup>24</sup>

I would thus like to suggest a solution to this question that Buchanan rejects, one he calls a “utilitarianism of rights,” according to which we should aim to resolve these questions by figuring out what would best promote self-determination for the most people (or best promote the average amount of self-determination among all people, or some other consequentialist sort of aggregation

17 Altman and Wellman, *A Liberal Theory of International Justice*.

18 Stilz, “The Value of Self-Determination.”

19 Nagel, “The Problem of Global Justice”; Banai, “Political Self-Determination and Global Egalitarianism.”

20 Levitov, “Human Rights, Self-Determination, and External Legitimacy.”

21 Moore, *A Political Theory of Territory*, 65.

22 Pevnick, *Immigration and the Constraints of Justice*.

23 Renzo suggests his theory is compatible at least with Moore and Stilz. He also cites three unpublished manuscripts that describe his own theory. Renzo, “Helping the Rebels,” 227, 232n28, 234n34.

24 I argue for this claim in Weltman, “Who Is the Self in Self-Determination?” and “Against Innovative Accounts of Self-Determination.”

scheme that works on the level of individuals), rather than being precious about making sure each particular group has its own right to self-determination.<sup>25</sup> Are there one or two groups in Cyprus? Pick the cleavage that would let the most people rule themselves rather than be ruled by others, or that would lead to the greatest amount of self-determination for any given individual that is compatible with at least that amount for everyone else, or some other individual-centric calculation.<sup>26</sup> One main worry with this utilitarianism of rights is that some groups would be thrown under the bus for the sake of other individuals in other groups. But there is no possibility of a world with perfect borders, especially if we are aiming to be practical and resolve the hardest cases, like the ones relating to intervention. There are always going to be clashes of self-determination with some winners and some losers. A utilitarianism of rights just aims to make sure we pick the losers by minimizing the number of people who lose out on self-determination (or the number of people with the minimal acceptable level of self-determination, or something like this), rather than via some other means that prioritize the choices of groups rather than outcomes for individuals.

We can picture examples where this gives us absurd results, like large State *A* intervening in small State *B* for the sake of promoting State *A*'s self-determination, leaving State *B* less self-determining. This is a possibility we open ourselves up to with any sort of utilitarian approach. In actual cases, though, we need not worry about these possibilities for the same reason a utilitarian of personal autonomy rights potentially does not need to worry about slavery. The utilitarian of personal autonomy rights argues that although, in principle, subjugating some might be a way of maximizing personal autonomy, in practice slavery is such a huge violation of autonomy that it is practically impossible to justify it with gains in personal autonomy for others. It is not important whether this argument vindicates utilitarianism of personal autonomy.<sup>27</sup> What is important is that the analogous claim is true with respect to political self-determination. This is because even in the rare cases in which the empirical claim is false in the political sphere, it is much less objectionable to compromise the autonomy of a group for the sake of other groups. As noted above, groups of people will always exist that are not able to be self-determining unless we compromise on some other group's self-determination, which means that we are going to have to

25 Buchanan, "The Ethics of Revolution and Its Implications for the Ethics of Intervention," 466. Recall that Buchanan is fine with weighing rights violations against each other—he just does not want to collapse entirely into a utilitarianism of rights.

26 This would likely entail the conclusion that there are two groups in Cyprus, but I do not want to commit to any particular solution to this actual case.

27 I think it does not. I thank an anonymous reviewer for suggesting that I address this worry.

compromise (whereas one might think we should never compromise in the case of individuals). In other words, political self-determination is not an inviolable right, at least for practical purposes, and if it is not inviolable, we might as well be utilitarians (of a sort).

So, we have vindicated Buchanan, albeit via a utilitarianism of rights that he wished to avoid. But we have chosen a utilitarianism of rights for groups in part because we are avoiding utilitarianism for individuals. For each individual person, what matters is that we aim to secure their belonging to a group in which they have self-determination, even if this requires seeing two groups where we might have seen one, or one where we before saw two. Exactly the reason utilitarianism is perhaps a bad fit for people—the fact that each person is a distinct, separate individual—is why utilitarianism is a fine fit for groups that exist for the sake of people. It is the people we care about, not the groups, and if some arrangement of groups would better serve people, we should endorse that arrangement. As Laura Valentini puts it, “individual human beings’ status as equal and ultimate units of moral concern” rules out using a collective agent, rather than individual agents, as the locus of concern: “the moral standing of a collective is explained by, and therefore conditional on, the collective’s serving the legitimate interests of individuals.”<sup>28</sup> We respect the rights of collectives only insofar as those collectives are good for individuals. This is incompatible with an inviolable group right to self-determination, but so much the worse for an inviolable group right to self-determination.

There is even a case to be made for pushing the logic of self-determination further. Charles Beitz argues that “claims of a right to self-determination, when pressed by or on behalf of residents of a colony, are properly understood as assertions that the granting of independence would help reduce social injustice in the colony,” and this point is not limited just to colonies.<sup>29</sup> “Self-determination,” he says, “is a means to the end of social justice,” not a fundamental right in and of itself.<sup>30</sup> Perhaps justice, broadly speaking, is what ultimately matters, and we ought to compromise self-determination for its sake.<sup>31</sup> This would move even further from Buchanan’s original goal, and leave us closer to a more traditional cosmopolitan, justice-centric view. This rejection of self-determination is “implicit in many cosmopolitan theories of global justice,” according to Lea Ypi.<sup>32</sup>

28 Valentini, “On the Distinctive Procedural Wrong of Colonialism,” 324.

29 Beitz, *Political Theory and International Relations*, 104.

30 Beitz, *Political Theory and International Relations*, 104.

31 I argue this in Weltman, “There Is Nothing *Per Se* Wrong with Colonialism.”

32 Ypi, “Territorial Rights and Exclusion,” 251.

Buchanan will see this as a downside, but for the reasons adduced above it may instead be a good result.

Whether we go this far, stop at a utilitarianism of rights, or even stop at acknowledging the differences between individuals and groups, the central point remains. Unless we can find some way around the worries about group agency raised above, it is not clear that Buchanan is wrong to posit a wider scope for intervention than Renzo supports.<sup>33</sup>

Ashoka University  
 danny.weltman@ashoka.edu.in

REFERENCES

- Altman, Andrew, and Christopher Heath Wellman. *A Liberal Theory of International Justice*. New York: Oxford University Press, 2009.
- Banai, Ayelet. "Political Self-Determination and Global Egalitarianism: Towards an Intermediate Position." *Social Theory and Practice* 39, no. 1 (January 2013): 46–69.
- Beitz, Charles. *Political Theory and International Relations*. Princeton: Princeton University Press, 1979.
- Buchanan, Allen. "The Ethics of Revolution and Its Implications for the Ethics of Intervention." *Philosophy and Public Affairs* 41, no. 4 (Fall 2013): 291–323.
- . *Justice, Legitimacy, and Self-Determination: Moral Foundations for International Law*. Oxford: Oxford University Press, 2004.
- . "Self-Determination, Revolution, and Intervention." *Ethics* 126, no. 2 (January 2016): 447–73.
- Jefferson, Thomas. Thomas Jefferson to Samuel Kercheval, July 12, 1816. In *The Thomas Jefferson Papers at the Library of Congress: Series 1: General Correspondence*. [https://www.loc.gov/resource/mtj1.049\\_0255\\_0262](https://www.loc.gov/resource/mtj1.049_0255_0262).
- Levitov, Alex. "Human Rights, Self-Determination, and External Legitimacy." *Politics, Philosophy and Economics* 14, no. 3 (August 2015): 291–315.
- Margalit, Avishai, and Joseph Raz. "National Self-Determination." *Journal of Philosophy* 87, no. 9 (September 1990): 439–61.
- Moore, Margaret. *A Political Theory of Territory*. New York: Oxford University Press, 2015.
- Nagel, Thomas. "The Problem of Global Justice." *Philosophy and Public Affairs* 33, no. 2 (March 2005): 113–47.

33 Thanks to an anonymous reviewer for this journal for their very helpful comments.

- Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.
- Pevnick, Ryan. *Immigration and the Constraints of Justice: Between Open Borders and Sovereignty*. Cambridge: Cambridge University Press, 2011.
- Renzo, Massimo. "Helping the Rebels." *Journal of Ethics and Social Philosophy* 13, no. 3 (July 2018): 222–39.
- Stilz, Anna. "The Value of Self-Determination." In *Oxford Studies in Political Philosophy*, vol. 2, edited by David Sobel, Peter Vallentyne, and Steven Wall, 98–127. New York: Oxford University Press, 2016.
- Walzer, Michael. *Just and Unjust Wars*. New York: Basic Books, 1977.
- Weltman, Daniel. "Against Innovative Accounts of Self-Determination." Unpublished manuscript.
- . "There Is Nothing *Per Se* Wrong with Colonialism." Unpublished manuscript.
- . "Who Is the Self in Self-Determination?" Unpublished manuscript.
- Valentini, Laura. "On the Distinctive Procedural Wrong of Colonialism." *Philosophy and Public Affairs* 43, no. 4 (Fall 2015): 312–31.
- Ypi, Lea. "Territorial Rights and Exclusion." *Philosophy Compass* 8, no. 3 (March 2013): 241–53.

## IS THERE VALUE IN KEEPING A PROMISE?

### A RESPONSE TO JOSEPH RAZ

*Crescente Molina*

SOME PHILOSOPHERS explain the demands for performance that promises impose on promisors by affirming that the fact of making a promise to  $\phi$  constitutes or creates a reason for  $\phi$ -ing. That means that this fact—making a valid promise—recommends or demands  $\phi$ -ing because there is value in someone’s  $\phi$ -ing if she has promised to do so. This view has famously been defended by Joseph Raz.<sup>1</sup> According to Raz, the fact of making a promise constitutes for the promisor both a reason to perform the promised act and a reason for not acting for at least some of the reasons that recommend something different than acting as promised (an “exclusionary” reason). On this view, if I promise to water your plants tomorrow I have both a reason to water your plants tomorrow and a reason for not acting for at least some of the considerations that recommend something different than watering your plants tomorrow (e.g., that my plants need more care than yours and it would be better if I use my time tomorrow for taking care of my plants rather than yours). Moreover, under this model, the reason to keep our promises is also a content-independent reason.<sup>2</sup> Leaving aside exceptional cases, we have a reason to perform our promises regardless of what the content of the promise is. If I promise to water your plants, I have a reason for watering your plants and for not acting for some of the reasons that militate against watering your plants, but it is the sole fact of this act (watering your plants) being *promised* that gives me those reasons and not other features or properties of it. As Raz does, I will call the content-independent reason to keep our promises a “promissory reason.”<sup>3</sup> I will argue that Raz’s account of the grounds of promissory reasons has some serious difficulties.

- 1 Raz, “Promises and Obligations,” 218–19, and “Is There a Reason to Keep a Promise?” 58. Arguably a similar position has been defended by Watson, “Promises, Reasons, and Normative Powers,” 155.
- 2 Raz, “Is There a Reason to Keep a Promise?” 59–63.
- 3 Raz, “Is There a Reason to Keep a Promise?” 59.

## 1. THE GROUNDS OF PROMISSORY REASONS

Different kinds of reasons demand that we keep our promises. For example, if I promise my best friend that I will go to her house for dinner on New Year's Eve, I have a reason to perform if I know that she will be disappointed if I do not show up, or if she prepared food for me that will be wasted if I do not show up, or even maybe simply because I enjoy being with her and it would be fun to go. All these reasons largely derive from the context in which the promise takes place, from the nature of the relationship between the promisor and the promisee, from the expectations created by the promise in the promisee, etc. Failing to keep a promise could imply nonconformity with all these different sorts of reasons. However, these reasons are not promissory reasons. As I previously defined them, promissory reasons are reasons to perform the promised act and to exclude some of the reasons that recommend otherwise that promises generate *qua* promises. But if promissory reasons obtain in virtue of facts that are not given by the nature of the promised act nor from all the different facts that may surround the making and performing of a promise (e.g., that the promisee or promisor will benefit from the performance), in virtue of which fact do we have a promissory reason?

According to some writers, the reason to keep our promises is grounded in a social practice of promise making and keeping. On this view (the "Practice View"), the practice of promising is a social rule by which—provided the requirements of the rule are fulfilled (e.g., that the promise was not given under duress)—promisors are obliged to perform the action they communicated to their promisee that they will perform. Writers like David Hume and John Rawls maintained that the rule of promising is valuable because it facilitates coordination, and since the possibility of coordination generated by the rule is realized only if a sufficient number of promisors perform their promises, the rule gives promisors a practice-based reason to keep their promises.<sup>4</sup> Against the Practice View, T. M. Scanlon has argued that we do not need something like the social practice of promising for us to be able to undertake promissory obligations. Scanlon holds that there is a general principle (the "principle of fidelity") that demands that, in the absence of justification, an agent who, with the aim of providing assurance to another person, voluntarily and intentionally causes that person to expect that he will perform or omit a certain action, and this person wants to be assured of this (and both parties know about the other's relevant beliefs and intentions), he should act as he said he would unless the other person consents otherwise. Making a promise activates this general principle of fidelity,

4 Hume, *A Treatise on Human Nature*, 331–37; Rawls, *A Theory of Justice*, 97–98, 303–8.

giving the promisor a reason to keep his promise.<sup>5</sup> Despite their deficiencies, the Practice View and Scanlon's account find a way of explaining why promises impose demands for performance on us.<sup>6</sup> They maintain that making a promise just triggers or activates normative reasons that, before making a promise, we had in a conditional form (i.e., reasons to protect and maintain valuable practices and reasons to avoid disappointing the expectations we cause in others).

Raz, on the other hand, refuses to adopt something like the Practice View or Scanlon's account, but nevertheless insists that the fact of making a valid promise creates or establishes a promissory reason.

## 2. RAZ ON PROMISSORY REASONS

Raz acknowledges that the fact that grounds promissory reason cannot be an interest or benefit of the promisee derived from the content of the promise itself, since, as we know, whether the promisee benefits from the performance of the promised act or not is a contingent matter that does not determine promises' bindingness.<sup>7</sup> In his early work on promises, Raz maintains that the ground for the promissory reason does not reside in some property of the promised act, but in the same fact or interest that justifies our power to make promises. In this early account, Raz maintains that what justifies our power to promise is our interest in creating and developing special bonds with other people.<sup>8</sup> Raz argues that the relationship between the content-independent reason to perform a specific promise (the promissory reason) and the reason or value that justifies our power to promise is a relation of practical inference.<sup>9</sup> Thus, if I promised to  $\phi$ , the proposition "I have a reason to  $\phi$  because I have promised to" is just the normative conclusion that follows from our interest in developing special bonds with others. However, Raz struggles to account for this conclusion. It is hard to see how our interest in being able to undertake special obligations with others gives us sufficient ground for affirming the existence of a promissory reason, that is, a content-independent reason to keep our promises. There may be value in

5 Scanlon, *What We Owe to Each Other*, 304.

6 For criticism of Scanlon's account, see Kolodny and Wallace, "Promises and Practices Revisited," 139–48; Owens, *Shaping the Normative Landscape*, ch. 9; Pratt, "Some Features of Promises and Their Obligations," 390–94. For criticism of the Practice View, see, e.g., Shiffrin, "Promising, Intimate Relationships, and Conventionalism," 519–23; Scanlon, *What We Owe to Each Other*, 309–11; Owens, *Shaping the Normative Landscape*, ch. 5.

7 Raz, "Promises and Obligations," 213–14.

8 Raz, "Promises and Obligations," 226–28.

9 Raz, "Promises and Obligations," 219. See also Raz, *The Morality of Freedom*, 84.

us being able to develop special obligations with others, but from this does not immediately follow that there must be value in us performing those obligations regardless of their content. It seems that we need some further argument to find the link between the consideration that justifies our ability to undertake voluntary obligations and the one that grounds promissory reasons.

In recent work Raz shares these doubts, and now tries to find something in each specific promise that, though different from a mere benefit to the promisee that arises from the content of the promised act, serves as a ground for promissory reasons.<sup>10</sup> Raz argues that the point of promising is that by making a promise the promisor gives or provides something for the promisee. He argues that in every valid promise what the promisor gives the promisee is the *normative assurance* that he will perform his promise. According to Raz, this normative assurance gives the promisor a right to performance and the power to waive such a right and release the promisor of his obligation.<sup>11</sup> However, it is still not very clear to me what Raz means when he affirms that promisors give their promisees the normative assurance of performance. As is clear, this assurance cannot be the, let us call it “epistemic” assurance, that the performance will actually happen. On many occasions promisees have no more reasons to *believe* that the promised act will take place just because the promisor promised it (the promisor might have a very bad promise-keeping record with the promisee, but his promise is nevertheless binding). Thus, the interest of promisees in each promise cannot be the interest in being more certain that the promised performance will occur.

Raz seems to argue that what the promisee receives from the promisor—that constitutes the content of the normative assurance that promisors provide to their promisees even when they do not benefit from the promised act—is an *opportunity* to develop interests in the performance.<sup>12</sup> A promisor, by making a specific promise, gives the promisee the option to develop interests in the performance, and even when the promisee does not benefit from the promise when it was made, he has the opportunity to become interested in the content of the promise, and it is such opportunity that the promisor must respect and protect by performing his promise if required by the promisee. Thus, the promisor has a reason to perform unless the promisee releases him.<sup>13</sup> But there are problems with this proposal.

Raz characterizes the fact in virtue of which promissory reasons obtain as being something that the promisor gives to the promisee by making the promise,

10 Raz, “Is There a Reason to Keep a Promise?” 67.

11 Raz, “Is There a Reason to Keep a Promise?” 71–72.

12 Raz, “Is There a Reason to Keep a Promise?” 73.

13 Raz, “Is There a Reason to Keep a Promise?” 77.

namely an opportunity to develop an interest in the performance. However, as I remarked before, what Raz refers to here must be something different than the mere expectations of performance that the promisor could generate in the promisee. So, if Raz's account claims to be something different than, for example, Scanlon's account, he must affirm that the opportunity to develop an interest in the performance the promisor gives to his promisee is generated regardless of whether or not the promisee forms any expectations that the promisor's performance will take place. However, if such an opportunity did exist, even when there are no expectations of performance, how would we characterize it?

One could argue that what the promisor gives to the promisee is authority over the promisor regarding the promised act. That is, the promisor gives to the promisee the power to control the permissibility of the promisor's failing to act as promised (i.e., the promisee acquires the power to cancel the promisor's obligation at will). David Owens defends this view, and maintains that, by making a promise, promisors specifically serve promisees' "authority-interest": that is, an interest in being able to control others' obligations. Promises would paradigmatically serve this interest even if they also served other interests of promisees (e.g., their interest in being able to predict promisors' future behavior).<sup>14</sup> Raz rejects this view. He states that he doubts the existence of Owens's authority-interest, and does not further engage with Owens's position.<sup>15</sup>

Thus, the problems for Raz's account persist. Until he gives us a clearer idea of what it is for promises to give normative assurance to promisees, and until he explains the sense in which this expands promisees' valuable opportunities, his account of the grounds of promissory reasons remains unconvincing.<sup>16</sup>

*University of Oxford*  
*crecente.molina@bnc.ox.ac.uk*

14 Owens, *Shaping the Normative Landscape*, chs. 4–6, and "A Simple Theory of Promising," 67–74.

15 Raz, "Is There a Reason to Keep a Promise?" 74.

16 Many thanks to James Edwards, Raffael Fasel, John Gardner, Manuel González, Felipe Jiménez, Christopher Kutz, Sebastian Lewis, David Owens, Olof Page, Alejandro Saenz, Sandy Steel, Samuel Williams, Tarek Yusari, an anonymous referee for the *Journal of Ethics and Social Philosophy*, and to the audience at the Oxford Jurisprudence Discussion Group for their helpful comments.

## REFERENCES

- Hume, David. *A Treatise on Human Nature*. Oxford: Oxford University Press, 2001.
- Kolodny, Niko, and R. Jay Wallace. "Promises and Practices Revisited." *Philosophy and Public Affairs* 31, no. 2 (Spring 2003): 119–54.
- Owens, David. *Shaping the Normative Landscape*. Oxford: Oxford University Press, 2012.
- . "A Simple Theory of Promising" *Philosophical Review* 115, no. 1 (January 2006): 51–77.
- Pratt, Michael. "Some Features of Promises and Their Obligations." *Southern Journal of Philosophy* 52, no. 3 (September 2014): 382–402.
- Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1999.
- Raz, Joseph. "Is There a Reason to Keep a Promise?" In *Philosophical Foundations of Contract Law*, edited by Gregory Klass, George Letsas, and Prince Saprai, 58–77. Oxford: Oxford University Press, 2014.
- . *The Morality of Freedom*. Oxford: Oxford University Press, 2009.
- . "Promises and Obligations." In *Law, Morality, and Society*, edited by P. M. S. Hacker and Joseph Raz, 210–28. Oxford: Oxford University Press, 1977.
- Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 2000.
- Shiffrin, Seana V. "Promising, Intimate Relationships, and Conventionalism." *Philosophical Review* 117, no. 4 (October 2008): 481–524.
- Watson, Gary. "Promises, Reasons, and Normative Powers." In *Reasons for Action*, edited by David Sobel and Steven Wall, 155–78. Cambridge: Cambridge University Press, 2011.