# ALLOWING THE WRONG, TO BRING ABOUT THE GOOD

#### David Faraci

Anne, a wealthy business owner, offers Bob, who is quite poor, a full-time position in her factory, where he will receive "substandard" compensation—compensation insufficient to meet his basic needs (food, shelter, etc.). Anne could offer Bob better compensation without lowering the net benefit of their interaction or eliminating her profits (though it might lessen them); she simply chooses not to. Bob correctly believes that factory work is better for him than any feasible alternative—including alternatives he would have if Anne's factory did not exist. He therefore consents to work in Anne's factory. No third parties are harmed by their interaction, and if Anne doesn't hire Bob, she won't hire anyone.

Many believe that in cases like this, Anne acts wrongly—specifically, that she wrongfully *exploits* Bob—despite her not being morally required to hire him in the first place. Various theories of exploitation have been deployed to vindicate this stance. For example, on one popular view—the one I assume herein for the sake of argument—Anne's offer is wrongfully exploitative because it is *unfair*.<sup>1</sup> Anne and Bob's employment relation creates a social surplus, and Bob's substandard compensation represents an unfairly small portion of that surplus.

Intriguingly, rather than directly attacking theories of wrongful exploitation, a number of critics argue that, fair or not, employment relations like Anne and Bob's cannot be wrongful because Anne consensually benefits Bob more than she was required to. According to what Alan Werthheimer calls "the nonworseness claim," a consensual, mutually beneficial transaction cannot be "morally worse" than its absence (Werthheimer 1999, 289–93). The apparent upshot for our

<sup>&</sup>lt;sup>1</sup> This is, among others, Werthheimer's (1999) view. Werthheimer introduces the nonworseness claim—my focus in Part I—as a potential objection to his own view.

example case is that if it is permissible for Anne not to hire Bob, it cannot be impermissible for her to provide him with exploitative compensation.

In this paper, I defend what I refer to as the "Reasonable View" against the nonworseness claim: Anne wrongfully exploits Bob, *even though* it is all-things-considered better that he be consensually benefitted than that he not be benefitted at all. I argue that once we understand the structure of the intuitive argument for the nonworseness claim, as well as the intuitions behind the Reasonable View, the former represents no threat to the latter. Accordingly, those who wish to argue that Anne does nothing wrong will have to directly engage with relevant theories of wrongful exploitation. This marks important dialectical progress.

As will become clear, it will be tempting given my arguments to conclude that only deontologists can embrace the Reasonable View.<sup>2</sup> I conclude by responding to a related objection: the Reasonable View should be rejected because it faces "the paradox of deontology." I close the main portion of the paper with a principled response to this objection. In an appendix, both as a secondary response to this objection, and to show that even consequentialists can embrace the Reasonable View, I develop a toy consequentialist theory consistent with that view.

Before moving on, some ground-clearing: Aside from pursuing the arguments discussed below, one might challenge the nonworseness claim by targeting its relevance, denying that cases like Anne and Bob's reflect what's going on in the real world of "sweatshop labor." For example, some critics argue that sweatshop employment offers are coercive, while some on the other side hold that offering greater compensation would harm third parties. If any such claims are true, actual

<sup>&</sup>lt;sup>2</sup> For our purposes, "deontology" is just non-consequentialism.

<sup>&</sup>lt;sup>3</sup> 'Sweatshop' here is just shorthand for (certain) employment arrangements with substandard compensation.

<sup>&</sup>lt;sup>4</sup> E.g., Barry and Reddy (2008) and Walzer (1983).

<sup>&</sup>lt;sup>5</sup> E.g., Powell (2014, chap. 3) and Zwolinski (2007). For a reply, see Kates (2015).

sweatshop employment may be importantly different from Anne and Bob's case. Such employment may be wrongful because it violates autonomy or permissible because paying more would be harmful on balance. These possibilities are irrelevant to the discussion on hand. Anne and Bob's case is described as it is because I am concerned specifically with the possibility of interactions that are wrongful, despite being morally optional, consensual and beneficial to all relevant parties.

### 1. The Nonworseness Claim and the Reasonable View

Consider two passages, the first from Werthheimer, introducing the nonworseness claim, the second from Matt Zwolinski, one of the nonworseness claims' champions:

Suppose that we criticize A for engaging in an unfair mutually advantageous transaction with B. A defends himself by appealing to what I shall call the *nonworseness claim*: "Look, you grant that I have a right not to transact with B. If B is better off if I engage in an unfair transaction with B than if I do not transact with B at all, it certainly can't be *morally worse* for me to engage in such a transaction with B. Given that I have a right not to transact with B and that transacting with B is not worse than not transacting with B, it can't be seriously wrong for me to engage in an unfair transaction with B." (Werthheimer 1999, 289)

[The nonworseness claim] holds that in cases where  $\mathcal{A}$  has a right not to transact with  $\mathcal{B}$ , and where transacting with  $\mathcal{B}$  is not worse for  $\mathcal{B}$  than not transacting with  $\mathcal{B}$  at all, then it cannot be seriously wrong for  $\mathcal{A}$  to engage in this transaction, even if its terms are judged to be unfair by some external standard . . . From a consequentialist moral framework, the nonworseness claim seems obviously true. (Zwolinski 2008, 357)<sup>6</sup>

Following these lines of thought, the sweatshop critic charges that Anne's hiring Bob is wrongfully exploitative because it is unfair. Anne replies that since Bob is better off than before, and she didn't have to benefit him at all, her benefitting him can't be wrong.

On its face, this is a rather strange reply. The critic can immediately respond that the presence of a benefit isn't what's at issue; *unfairness* is. Zwolinski's contention that "the

<sup>&</sup>lt;sup>6</sup> This quotation is from a paper about price-gouging. Zwolinski also deploys the nonworseness claim in his (2007) discussion of sweatshops; but his comments in the price-gouging paper are, I think, more illuminating.

nonworseness claim seems obviously true" for consequentialists further highlights this oddness. After all, Werthheimer and Zwolinski both *just told us* what moral value Anne's exploiting Bob might negatively impact: fairness. Nothing prevents consequentialists from acknowledging the value of fairness. And so nothing prevents consequentialists from claiming that even though "transacting with B is not worse for B than not transacting with B at all," (bolding mine) it is nevertheless worse all things considered because of its unfairness.

As an analogy, suppose someone were to propose a version of the nonworseness claim that lacks the consent condition:

**NWC-WELFARIST** The state of affairs in which A interacts with B cannot be all-things-considered worse than the state of affairs in which A does not interact with B if: (1) the interaction is better for B than non-interaction and (2) such interaction has no negative effects on others.

It would be very odd for someone to claim that NWC-WELFARIST is obviously true for consequentialists. It is only obviously true for consequentialists who deny the value of anything other than welfare—i.e., for welfarists. Likewise, the nonworseness claim is only obviously true for consequentialists who deny the value of anything other than welfare and autonomy (or whatever ungirds the importance of consent). And there are plenty outside of that camp.

If the nonworseness claim isn't baldly question-begging, something else must be going on here. The answer, I submit, is that both Werthheimer and Zwolinski expect that their readers will share an *evaluative* intuition: the state of affairs in which Bob is consensually benefitted, even unfairly, is all-things-considered better than the state of affairs in which he is (fairly) left in poverty. This isn't just a claim about welfare, about what's good *for Bob*; it's a claim about what's best with respect to what matters morally, and thus, implicitly, a claim that welfare matters more than fairness (at least here).

This thought is perhaps best motivated by considering our obligations as third parties—something the literature often mentions. Assuming we can't force Anne to hire Bob, the most we as

third parties can do is ensure that *if* she hires him, she offers him better compensation—say, by instituting a minimum wage. Suppose we are confident that this would lead Anne to not hire Bob. Bob is already so poorly off; surely we shouldn't do anything that will eliminate his best option, even to ensure that things are fair. Instead, it seems we should allow Anne to hire Bob, regardless of his compensation's fairness, so as to bring about more good. With this in mind, the implicit argument that it can't be wrong for Anne to exploit Bob is as follows:

**NWC-EVALUATIVE** The state of affairs in which A engages in interaction N with B cannot be all-things-considered worse than the state of affairs in which A does not interact with B if: (1) N is better for B than non-interaction; (2) B consents to N; and (3) N has no negative effects on others. Therefore...

**NWC-DEONTIC** ...it cannot be impermissible for A to engage in N with B if: (1) it is permissible for A to not interact with B; (2) N is better for B than non-interaction; (3) B consents to N; and (4) N has no negative effects on others.

If we read the argument this way, the passages above make much more sense. For Werthheimer, Anne is not begging the question because she assumes her opponent will agree that it's better that Bob be treated unfairly than left in poverty. Zwolinski doesn't really mean that no consequentialist can deny the nonworseness claim; he means that no consequentialist who accepts NWC-EVALUATIVE can deny NWC-DEONTIC. After all, for consequentialists, more goodness means more rightness.

With this version of the argument in hand, it is easy to see how the nonworseness claim is meant to threaten the Reasonable View, taken as the conjunction of four claims:

**PRIOR** Anne has no prior moral obligation to hire Bob nor to benefit him to the extent she does in hiring him with substandard compensation.

**ALLOWED** It is morally impermissible for third parties to take actions expected to lead Anne not to hire Bob, because...

**BETTER** ... the state of affairs in which Anne employs Bob with substandard compensation is all-things-considered better than the state of affairs in which Anne does not employ Bob at all.

**UNFAIR** It is morally impermissible for Anne to hire Bob with substandard compensation because her doing so would be unfair.

Given the details of Anne and Bob's case, the conjunction of PRIOR and NWC-DEONTIC entails the negation of UNFAIR. The Reasonable View is therefore inconsistent with the nonworseness claim, taken as the conjunction of NWC-EVALUATIVE and NWC-DEONTIC (or just the latter). In order to resist the nonworseness claim, the champion of the Reasonable View must either deny NWC-EVALUATIVE (a strategy I won't pursue) or deny that it implies NWC-DEONTIC. Taking the latter tack requires rejecting the following general constraint on the structure of moral theories:

**SYMMETRY** The *evaluative* ordering of states of affairs must match the *deontic* ordering of actions that (are expected to) bring about those states of affairs.

The Reasonable View seems inconsistent with SYMMETRY, for it entails that one impermissible action—Anne's exploiting Bob—generates greater expected value than one permissible action—Anne's not hiring Bob at all (though in the appendix I demonstrate that the Reasonable View is actually consistent with SYMMETRY).

The obvious claim here is that while the nonworseness claim might provide a good argument against *consequentialist* versions of the Reasonable View, it has no force against *deontological* ones. After all, rejection of SYMMETRY is a hallmark of deontology. Consider Thomson's (1985) classic case of the surgeon who considers sacrificing one patient to save five others, often deployed as a counterexample to consequentialism. The deontologist's position is that an impermissible action—killing the one to save the five—generates greater expected value than a permissible one—allowing the five to die.

Nevertheless, Zwolinski maintains that the nonworseness claim "is meant to have traction for deontological theories as well," because "it is puzzling how it could be worse by any moral standard  $\dots$  for  $\mathcal A$  to provide some help than it is for him to provide none." Following Wertheimer,

Zwolinski points out that to get around such puzzlement, deontologists would need to defend an "interaction principle" according to which "there are special moral constraints that apply to our interactions with others *if we choose to interact with them,* but which do not have anything to say about those who choose not to interact at all" (Zwolinski 2008, 357).

Zwolinski finds such interaction principles implausible. Indeed, even Wertheimer, who defends them, proceeds with some trepidation. The reason, I suspect, is that it is tempting to think of interaction principles as counterintuitively placing extra moral burdens on the virtuous, just as Zwolinksi charges:

[T]heories which subject A to a moral obligation to rescue B on condition of their interaction seem to suffer from both a defect of unfairness and a serious internal tension. A's interaction with B is supererogatory, done with B's consent, violates no independent moral constraints, and benefits B. Why, then, should the interaction itself place A under new moral obligations toward B, beyond those to which A and B mutually agree? On the face of it, it seems unfair to burden A with this extra requirement given that he is *already* doing more than is morally required of him. (Zwolinski 2008, 359)

To make this concrete, consider an interaction principle that requires you to donate 10% of your income to charity, but only if you choose to donate anything at all. This does seem unfair; indeed, it seems downright bizarre. Presumably, something like the moral significance of welfare undergirds the fact (if it is a fact) that donating to charity is supererogatory. But if the moral significance of welfare merely *licenses* donating, why would it *require* donating more once you begin? This seems no less bizarre as a deontological side constraint than it would as an implication of a consequentialist calculus.<sup>7</sup>

That bizarreness, however, is not a feature of interaction principles per se; it is a feature of *monistic* interaction principles on which a single morally significant feature requires more of one as one responds properly to it. Crucially, this is not the sort of interaction principle at work in the

<sup>&</sup>lt;sup>7</sup> Though for an attempt to make this more palatable, see Snyder (2008).

Reasonable View. The intuition behind UNFAIR is not that Anne has to benefit Bob *more* because she has benefited him *some*. Rather, it is that she must benefit him more because she is keeping for herself an *unfair* proportion of the social surplus created by their interaction. Anne's hiring Bob generates new obligations not because the moral significance of welfare requires more of her now that she's helped him, but because the moral significance of *fairness* has become salient within their interaction.<sup>8</sup> Once we understand this, the worry that interaction principles place unfair burdens on the generous vanishes. Anne falls afoul of her interaction obligation only when she seeks to take too much for herself.<sup>9</sup>

This is the crucial feature of the Reasonable View that the above criticisms miss. It is misleading to say that Anne is "providing help." If that's all she were doing, then of course it would be better than doing nothing. But her interaction with Bob is mutually beneficial, and fairness may well have something to say about how much she must distribute those benefits. This is simply a reflection of the fact that some values may generate new obligations within—perhaps only within—interactions. Given this, we can vindicate the Reasonable View with the following, perfectly sensible deontological view. Welfare is a good-making feature of states of affairs (as per BETTER), but promoting the good is merely supererogatory, so Anne is not obligated to hire Bob (as per PRIOR). There are also side constraints: it is impermissible to distribute benefits unfairly (as per UNFAIR) or to force people to do what they are morally permitted not to do (as per ALLOWED).

<sup>&</sup>lt;sup>8</sup> This is assuming that the distribution of resources was fair before their interaction, or at least that its prior unfairness had no deontic implications for Anne.

<sup>&</sup>lt;sup>9</sup> Actually, if she were to pay Bob *more* than was fair, that might violate the obligation as well. In that sense, perhaps such an obligation would have problematic implications for the generous. But the problem wouldn't be that they are unfairly burdened; it would be that we failed to accurately accommodate certain relevant intuitions, perhaps about the permissibility of waiving one's right to one's due, or the priority we should give to helping the worse off, even at the expense of fairness.

## 2. The Paradox of Deontology

Given the above, it might seem *only* deontologists can embrace the Reasonable View. Some, I expect, would find this a disappointing result. But Benjamin Ferguson takes an even dimmer view, arguing that any theory that violates SYMMETRY, like the Reasonable View, should be rejected because it faces a dilemma: "it is either 'axiologically incongruous' or encounters the 'paradox of deontology'." On the first horn, evaluative properties and deontic requirements are independent of one another, leaving us with a "moral framework [that] suffers from a kind of 'two-mindedness' about moral value." On the second horn, the theory claims to ground deontic commitments in evaluative ones, yet includes principles requiring one to bring about *less* of what's valuable. Ferguson charges that "[p]rinciples that prohibit what they are designed to promote suffer from a form of inconsistency" (Ferguson 2011, 14).

I suspect these complaints seem damning here (to those who find them so) for the same reason the nonworseness claim does. For many, the intuition that it would be better for Anne to exploit Bob than to leave him in poverty survives reflection, even in the face of the intuition that Anne is doing something impermissible. By contrast, the typical consequentialist response to Thomson's surgeon case is not to raise Ferguson's dilemma, but rather to accept the deontologists' deontic claim but deny their evaluative one, agreeing that it is wrong to carve up the one to save the

<sup>&</sup>lt;sup>10</sup> Ferguson's solution is to accept a more limited version of the nonworseness claim that would allow us to avoid axiological asymmetry by rejecting BETTER: "If one or more strictly Pareto improving and consensual transactions are possible, then for at least one of these transactions, transacting is morally better than not transacting" (Ferguson 2011, 15). I think this fails to sufficiently capture the intuitiveness of NWC-EVALUATIVE and BETTER, including their plausibility as grounds for accepting ALLOWED.

five, but maintaining that this is because the state of affairs in which the one dies is not all-thingsconsidered better after all.

Here again it can be useful to compare judgements about third-party interference. Most would judge not only that it would be wrong for the surgeon to kill the one to save the five, but also that third parties should prevent the surgeon from doing so. As some might put it, it is not merely wrong for the surgeon to violate the patient's right to life (supposing that's what makes it wrong); it is in some important sense agent-neutrally bad for the state of affairs in which the one dies to save the five to be brought about. It is not hard to see how this could bolster consequentialists' willingness to maintain that the latter is not really a better state of affairs at all.

The upshot is that those with the relevant intuitions about Anne and Bob's case seem destined to deny SYMMETRY, whereas most who think it wrong to carve up the one to save the five are not. Highlighting this—along with the "agential asymmetry" regarding what Anne should do vs. what third parties should allow—may well exacerbate the appearance of tension between our evaluative and deontic commitments. This, in turn, may lend force to worries like Ferguson's, just as it did the nonworseness claim.

Nevertheless, Ferguson's argument fails because it is question-begging. He seems absolutely right that the correct deontic principles should reflect what matters morally. But he simply assumes that *features of states of affairs* exhaust what matters morally. That is just to assume that consequentialism is true. Suppose instead that *acting fairly* itself matters morally. It may be that this cannot be, or is not *best*, <sup>11</sup> understood as a reflection of the value of some feature of states of

<sup>&</sup>lt;sup>11</sup> This may all be moot if every moral theory can be "consequentialized." Campbell Brown (2011) argues that they cannot. (He references Dreier (1993), Louise (2004) and Portmore (2007) as theorists who endorse the thesis that they can "or something close to it" (Brown 2011, 750, fn.1).) However, Brown defines consequentialism in terms of "maximizing the good," which some might regard as too narrow. In consequentializing the Reasonable View in the appendix, I take consequentialism to include any theory whose

affairs—e.g., the idea is not that we should maximize fair actions. That does not reveal some deep structural problem with theories that endorse the initial claim; it reveals only that the view is not a form of consequentialism.<sup>12</sup>

### 3. Conclusion

Some may wish to challenge the nonworseness claim's implications for the sweatshop debate by insisting that employers are antecedently obligated to benefit their employees (contra PRIOR) or that the state of affairs in which someone is unfairly benefitted is worse than the one in which he is not benefitted at all (contra BETTER). But those who think employers like Anne are doing something wrong need not reject these claims. They need only embrace the possibility that some other value,

deontic commitments are entailed by its evaluative commitments, where value attaches to features of states of affairs. (I take no position here on whether all views can be consequentialized on this understanding of consequentialism. My suspicion is that they can, but that many—including, arguably, the Reasonable View—are not *best* understood as forms of consequentialism.)

This is perhaps easiest to see on—but crucially does not *demand*—a buck-passing view on which facts about value are a function of facts about reasons, and (say) one always has decisive reason to distribute social surpluses fairly. Note that this is not just a problem for Ferguson, but for many (perhaps all) appeals to the "paradox of deontology." For example, Alexander and Moore (2015) write that "[d]eontologists need their own, non-consequentialist model of rationality, one that is a viable alternative to the intuitively plausible, 'act-to-produce-the-best-consequences' model of [moral] rationality that motivates consequentialist theories. Until this is done, deontology will always be paradoxical." But while deontologists do need a model of rationality, it is a mistake to claim that without one, the view is "paradoxical." That is to question-beggingly judge deontological theories' coherence against a background assumption that acting morally is about producing good outcomes. (This is especially problematic given that deontologists are often working "bottom-up," starting from certain *substantive* intuitions about what morality or rationality requires, ones the presumed consequentialist default can't (easily) accommodate.) All of this is symptomatic of the all-too-common, illegitimate assumption that to say that something matters morally (is valuable) is to say that it merits our promoting its obtaining as a feature of states of affairs. For a relevant, useful distinction between "Moorean" and "Kantian" value concepts see Bradley (2006).

like fairness, makes it impermissible to exploit even while consensually benefitting. Nothing prevents their doing this while also accepting the evaluative claim that the state of affairs in which someone is consensually benefitted is all-things-considered no worse or better than the one in which he is not (as per NWC-EVALUATIVE and BETTER). Finally, given that it is an open question whether evaluative claims about features of states of affairs exhaust claims about what matters morally, there is no principled objection to the structure of the Reasonable View. I therefore conclude that the nonworseness claim, and its intuitive (evaluative) roots, pose no threat to the Reasonable View, at least as a commitment of a deontological theory. It may be that we should sometimes allow the wrong, to bring about the good.

### Appendix: A Consequentialist Version of the Reasonable View

In this appendix, I show that even if one assumes that features of states of affairs exhaust what matters morally, one need not reject the Reasonable View. I develop a toy consequentialist theory, "Toy-leology," that is compatible with the Reasonable View (given two further stipulations about Anne and Bob's case, explicated below). Toy-leology is consequentialist in that all of its deontic commitments are entailed by its evaluative ones, which all concern features of states of affairs. Toy-leology is also consistent with SYMMETRY.

Some, of course, endorse narrower definitions of consequentialism. But note that it is already doubtful that we can rest easy with a traditional "maximize the good" understanding of consequentialism, since it is a point of agreement in the relevant dialectic that benefitting Bob is supererogatory. In any case, whether or not Toy-leology appeals to consequentialists, it avoids structural objections like Ferguson's.<sup>13</sup>

<sup>&</sup>lt;sup>13</sup> Thanks to Daniel Nolan for discussion on this point.

Toy-leology recognizes the values of autonomy, welfare and fairness. ALLOWED suggests that autonomy (which I'm taking to undergird the impermissibility of forcing people to do what they are morally permitted not to do) is of the utmost importance. For simplicity's sake, Toy-leology takes autonomy's value to be lexically prior to the values of both welfare and fairness.

To simplify things going forward, two further stipulations about Anne and Bob's case: first, Anne's exploiting Bob would deplete net fairness; second, forcing Anne to hire Bob would deplete net autonomy. 14 Now consider the following states of affairs:

- **S**<sub>1</sub>. The result of Anne's choosing to exploit Bob.
- $S_2$ . The result of Anne's choosing to not hire Bob.
- S<sub>3</sub>. The result of Anne's choosing to hire Bob and compensate him fairly.
- **S<sub>4</sub>.** The result of Anne's being forced to hire Bob and compensate him fairly.

BETTER suggests that welfare sometimes outweighs fairness, that  $S_1 > S_2$ . But because of SYMMETRY, UNFAIR suggests that the reverse can also be the case, that  $S_2 > S_1$ . Toy-leology remains consistent by taking the weightings of welfare and fairness to be *conditional*. To see the relevant conditions, notice that the states of affairs accessible to Anne (through action or inaction) are not the same as those accessible to third parties. Anne can access  $\{S_1, S_2, S_3\}$  but not  $S_4$ . Third parties can access  $\{S_1, S_2, S_4\}$  but not  $S_3$ . What's more, there is a principled difference between these sets: only the latter involves *value conflict*, i.e., lacks a member that weakly dominates with respect to all values.

<sup>&</sup>lt;sup>14</sup> The first stipulation is fairly natural. It *could* turn out that in exploiting someone, we makes things more fair overall. But given Bob's poverty, it does not seem much of a stretch to suppose that the local unfairness of Anne's exploitative offer would lead to a *globally* less fair outcome. The second stipulation might seem more problematic. But note that *any* consequentialist view that accepts PRIOR and ALLOWED—including those of Zwolinski and other champions of the nonworseness claim—will have to somehow vindicate a fairly general prohibition on forcing people to do supererogatory things.

Start with the set of states accessible to third parties:  $\{S_1, S_2, S_4\}$ . Suppose  $S_2$  contains x welfare, y fairness and z autonomy. Third parties can access this state by (say) instituting a minimum wage and leading Anne to *never* hire Bob. By doing nothing, third parties can instead access  $S_1$ , which contains >x welfare, <y fairness and  $\ge z$  autonomy. Or by forcing Anne to hire Bob, third parties can access  $S_4$ , which contains >x welfare,  $\ge y$  fairness and <z autonomy. Within this set, there is value conflict: no member has at least as much welfare, fairness and autonomy as the others. Something must be sacrificed.

By contrast, consider the set of states accessible to Anne:  $\{S_1, S_2, S_3\}$ . Again,  $S_2$  contains x welfare, y fairness and z autonomy. By exploiting Bob, Anne can instead access  $S_1$ , which contains >x welfare, <y fairness and  $\ge z$  autonomy. Or by choosing to hire Bob and compensate him fairly, Anne can access  $S_3$ , which contains >x welfare,  $\ge y$  fairness and  $\ge z$  autonomy. Here, there is no conflict. Anne can promote welfare without sacrificing anything of value, by choosing to hire Bob and compensate him fairly  $(S_3)$ .

Following suit, Toy-leology holds that (a) welfare's value is lexically prior to fairness' conditional on the *presence* of value conflict, while (b) fairness' value is lexically prior to welfare's conditional on the *absence* of value conflict. Holding autonomy fixed, this is as follows:

**WEIGHTING** Given a set of states of affairs  $\{S_1-S_n\}$ ,  $S_1$  is all-things-considered better than  $S_2$  if and only if (i)  $S_1$  contains more fairness but no less welfare than  $S_2$  or (ii)  $S_1$  contains more welfare but no less fairness than  $S_2$  or (iii)  $S_1$  contains more welfare but less fairness than  $S_2$  and there is no  $S_X$  with no less welfare than  $S_1$  and no less fairness than  $S_2$  or (iv)  $S_1$  contains more fairness but less welfare than  $S_2$  and there is some  $S_X$  with no less welfare than  $S_1$  and no less fairness than  $S_2$ .

WEIGHTING is ad hoc, but not unmotivated. It serves to capture, within a consequentialist framework, an intuition about how the value of fairness constrains us—an intuition typically associated with deontology. Welfare, the thought goes, matters enough that when one must choose between welfare and fairness, one should choose welfare. This is why, where welfare and fairness conflict, the former has priority. But fairness also matters. And when we are promoting welfare, we

are required to distribute the surplus as fairly as we can. This is why, where welfare and fairness do not conflict, one may promote the former only insofar as one avoids depleting the latter.

For determining the truth value of comparative evaluative claims, Toy-leology takes the relevant set of states to be all and only those mentioned in the relevant claim. Thus, Toy-leology reads BETTER as  $S_1 > S_2 \mid \{S_1, S_2\}$ . The set  $\{S_1, S_2\}$  involves value conflict: the state resulting from Anne's exploiting Bob  $(S_1)$  has more welfare, but less fairness, than the state resulting from her not hiring him  $(S_2)$ . Welfare's priority therefore applies and Toy-leology entails BETTER.

For determining the truth value of deontic claims about an action  $\varphi$  by agent G, Toy-leology takes the relevant set of states of affairs to be all and only those states (expected to be) accessible to G through action, plus the counterfactual status quo  $S_C^G$ , the state resulting from G's inaction. Were Toy-leology a form traditional, maximizing consequentialism, it would hold that G is obligated to take whatever action would bring about the state of affairs with the greatest (expected) net value. However, as per PRIOR, Toy-leology needs to entail that though it would be good for Anne to benefit Bob (as per BETTER), she isn't *obligated* to (as per PRIOR).

Consequentialist theories have various tools for accommodating supererogation. One popular option is to develop the theory as a form of *satisficing* consequentialism. So as to avoid questions about the specific amounts of welfare and fairness in Anne and Bob's case, Toy-leology sets the satisficing point at the counterfactual status quo. All value-*promoting* actions are supererogatory, while all value-*depleting* actions are impermissible, as per:

**PROMOTE** To *promote* value is to take an action  $\varphi$  such that the state of affairs that results from  $\varphi$ -ing has greater expected value than the state of affairs that would result from not  $\varphi$ -ing.

**DEPLETE** To *deplete* value is to take an action  $\varphi$  such that the state of affairs that results from  $\varphi$ -ing has less expected value than the state of affairs that would result from not  $\varphi$ -ing.<sup>15</sup>

Again, the set of states accessible to Anne is  $\{S_1, S_2, S_3\}$ .  $S_C^A$  is  $S_2$ . Since value-promoting actions are not required, it is permissible for Anne to take no action and remain in  $S_2$ , as per PRIOR. In exploiting Bob, Anne transitions from  $S_2$  to  $S_1$ .  $S_1$  has less fairness than  $S_2$ . Since all value-depleting actions are impermissible,  $(S_C^A \to S_1)$  is impermissible, as per UNFAIR. In hiring Bob and compensating him fairly, Anne transitions from  $S_2$  to  $S_3$ .  $S_3$  has more welfare than  $S_2$  and no less fairness or autonomy. Since all value-promoting actions are supererogatory,  $(S_C^A \to S_3)$  is supererogatory. Following the natural thought that supererogatory actions are "deontically superior" to merely permissible ones, it follows that  $(S_C^A \to S_3) > (S_C^A \to S_2) > (S_C^A \to S_1)$ . Since Toy-leology entails  $S_3 > S_2 > S_1 \mid \{S_1, S_2, S_3\}$ , this is consistent with SYMMETRY.

For third parties, the set of accessible states is  $\{S_1, S_2, S_4\}$ .  $S_C^{3P}$  is  $S_1$ , since if we do nothing, Anne will exploit Bob. Since value-promoting actions are not required, it is permissible for third parties to do nothing and remain in  $S_1$ . In instituting a minimum wage or forcing Anne to hire Bob, we transition from  $S_1$  to  $S_2$  or  $S_4$  (respectively), which depletes welfare or autonomy (respectively), so  $(S_C^{3P} \to S_2)$  and  $(S_C^{3P} \to S_4)$  are both impermissible, as per Allowed. Following the natural thought that given autonomy's lexical priority, autonomy-depleting actions are "deontically inferior" to welfare-depleting ones, it follows that for third parties,  $(S_C^{3P} \to S_1) > (S_C^{3P} \to S_2) > (S_C^{3P} \to S_4)$ . Since Toy-leology entails  $S_1 > S_2 > S_4 \mid \{S_1, S_2, S_4\}$ , this, too, is consistent with SYMMETRY.

<sup>15</sup> Neither inaction nor negatively defined actions such as 'not giving Bob a million dollars' are actions here (i.e., they cannot be the value of  $\varphi$ ). If  $\varphi$  were inaction, determining the counterfactual status quo would require determining what action Anne would have taken if she took any action. If  $\varphi$  were 'not giving Bob a million dollars', Anne would be required to give Bob a million dollars, since it is impermissible to deplete welfare, and Anne's not giving Bob a million dollars depletes welfare compared to the state of affairs resulting from her refraining from not giving Bob a million dollars.

Summing up: Promoting welfare is not required, and thus it is permissible for Anne not to hire Bob, as per PRIOR. Anne's compensating Bob depletes fairness (and therefore, given that she faces no value conflict, overall value) and is therefore impermissible, as per UNFAIR. The state of affairs in which Anne exploits Bob contains more welfare but less fairness than the state of affairs in which she doesn't hire him at all. Given welfare's lexical priority under value conflict (which there is between these two states), it follows that the former state of affairs is all-things-considered better, as per BETTER. Finally, forcing Anne to hire Bob decreases autonomy (and therefore, given autonomy's unconditional lexical priority, overall value) and is therefore impermissible, as per ALLOWED.

Toy-leology entails and therefore serves as a possibility proof for consequentialist endorsement of the Reasonable View. It also avoids both horns of Ferguson's dilemma. Because the view's deontic commitments are a direct consequence of its evaluative ones, it is neither objectionably "two-minded," nor does it recommend bringing about less of what it deems valuable.

# Acknowledgements

Thanks to Michael Kates for work and discussions that inspired this paper. Thanks to Tristram McPherson for helpful comments. And thanks to John Hasnas, Peter Jaworski, Govind Persad, Abe Singer and other participants in the Georgetown Institute for the Study of Markets and Ethics Faculty Workshop Series.

### References

Alexander, L., and M. Moore. 2015. "Deontological Ethics." In *The Stanford Encyclopedia of Philosophy*, edited by E.N. Zalta, Spring 2015.

http://plato.stanford.edu/archives/spr2015/entries/ethics-deontological/.

Barry, C., and S. Reddy. 2008. *International Trade and Labor Standards: A Proposal for Linkage*. Columbia University Press.

Bradley, B. 2006. "Two Concepts of Intrinsic Value." *Ethical Theory and Moral Practice* 9 (2): 111–30. Brown, C. 2011. "Consequentialize This." *Ethics* 121 (4): 749–71.

Dreier, J. 1993. "Structures of Normative Theories." Monist 76 (1): 22-40.

Ferguson, B. 2011. "The Paradox of Exploitation." Erkenntnis, 1–22.

Kates, M. 2015. "The Ethics of Sweatshops and the Limits of Choice." *Business Ethics Quarterly* 25 (2): 191–212.

Louise, J. 2004. "Relativity of Value and the Consequentialist Umbrella." *Philosophical Quarterly* 54 (217): 518–36.

Portmore, D.W. 2007. "Consequentializing Moral Theories." *Pacific Philosophical Quarterly* 88 (1): 39–73.

Powell, B. 2014. Out of Poverty: Sweatshops in the Global Economy. New York: Cambridge University Press.

Snyder, J.C. 2008. "Needs Exploitation." Ethical Theory and Moral Practice 11 (4): 389–405.

Thomson, J.J. 1985. "The Trolley Problem." Yale Law Journal 94 (6): 1395–1415.

Walzer, M. 1983. Spheres of Justice: A Defense of Pluralism and Equality. New York: Basic Books.

Werthheimer, A. 1999. Exploitation. Revised edition. Princeton University Press.

Zwolinski, M. 2007. "Sweatshops, Choice, and Exploitation." *Business Ethics Quarterly* 17 (4): 689–727.

——. 2008. "The Ethics of Price Gouging." Business Ethics Quarterly 18 (3): 347–78.