Consequentializing Constraints

A Kantsequentialist Approach

Douglas W. Portmore

Abstract: There is, on a given moral view, a constraint against performing acts of a certain type if that view prohibits agents from performing an instance of that act-type even to prevent two or more others from each performing a morally comparable instance of that act-type. The fact that commonsense morality includes constraints has been seen by several philosophers as a decisive objection to act-consequentialism. Despite this, I argue that constraints are more plausibly accommodated within an act-consequentialist framework than within the more standard side-constraint framework. For I argue that when we combine agent-relative act-consequentialism with a Kantian theory of value, we arrive at a version of consequentialism—namely, *Kantsequentialism*—that has several advantages over the standard side-constraint approach to accommodating constraints. What's more, this version of consequentialism avoids the disadvantages that critics of consequentializing have presumed that any such theory must have.

1. Constraints

There is, on a given moral view, a constraint against performing acts of a certain type if that view prohibits agents from performing an instance of that act-type even to prevent two or more others from each performing a morally comparable instance of that act-type.[1] Thus, there is, on

[1] As Scheffler puts it, "an agent-centred restriction [or constraint] is, roughly, a restriction which it is at least sometimes impermissible to violate in circumstances where a violation would serve to minimize total overall violations of the very same restriction, and would have no other morally relevant consequences" (1985, 409). Note, then, that a constraint is not simply a prohibition against performing an act of a certain type even to prevent two or more others from each performing an instance of that act type. After all, classical utilitarianism would prohibit you from failing to maximize utility even to prevent two others from each failing to maximize utility. But, in any such case, your act couldn't be morally comparable to those of the other two. For your act would fail to maximize utility only if it resulted in a net loss of utility that's greater than the combined net loss resulting from those of the other two. So, despite what some have claimed (e.g., Ridge 2009, 422), classical utilitarianism doesn't imply that there is a constraint against failing to maximize utility or any other type of act, for it doesn't prohibit you from performing an act-type even to prevent two others from each performing a *morally comparable* instance of that act-type.

commonsense morality, a constraint against breaking a promise given that it prohibits agents from breaking a promise even to prevent two others from each breaking a morally comparable promise. The fact that commonsense morality includes such a constraint shows that it doesn't simply take promise-breakings to be something bad that agents should, other things being equal, minimize. Additionally, commonsense morality must hold either that agents are prohibited from breaking a promise in the pursuit of their ends or that agents are required both to adopt some end that would be achieved only by their not breaking their promises and to give this end priority over that of their minimizing promise-breakings overall.

I'll argue that it's the latter. For I'll argue that it has at least three advantages and no disadvantages. Admittedly, some philosophers have argued that it has several disadvantages, but I'll show that they're mistaken. For I'll show both that we can adopt the latter by combining agent-relative act-consequentialism with a Kantian theory of value and that, when we do, we avoid these putative disadvantages while holding on to all the advantages. Thus, if my arguments succeed, they'll demonstrate that constraints are more plausibly accommodated within an act-consequentialist framework.

2. Two Opposing Approaches to Accommodating Constraints

As we've just seen, there are two competing approaches to accommodating constraints. One approach is to include an outright restriction on the types of acts that agents are permitted to perform in the pursuit of their ends. The other approach is to give agents an end that would be achieved only by their not performing acts of certain types and to insist that this end has priority over that of their minimizing overall performances of these act-types. The former is what I'll call *the side-constraint approach*, and the latter is what I'll call *the teleological approach*. The key difference is that whereas the teleological approach allows that an agent's other legitimate ends could—at least, in principle—affect the permissibility of their infringing upon a constraint, the side-constraint approach doesn't. Note, then, that any constraint-accepting theorist must be either a teleologist or a side-constraint theorist. For any theorist who accepts that there is a constraint against, say, φ-ing must either accept or deny that the permissibility of φ-ing

depends on what the agent could achieve by φ-ing and how that compares to what they could achieve by refraining from φ-ing. If they accept this, they're a teleologist. And if they deny this and, so, accommodate the constraint against φ-ing by holding that agents are prohibited from φ-ing for the sake of any end, then they're a side-constraint theorist. Also, note that, like Judith Jarvis Thomson, I'll distinguish between merely *infringing upon* a constraint and *violating* a constraint. A subject infringes upon a constraint against their performing a given act-type if and only if they perform an instance of that act-type. By contrast, a subject violates that constraint if and only if they not only infringe upon it but also act wrongly in doing so (Thomson 1986, 51). Thus, not all infringements are wrong, but those that are wrong are called *violations*.

2.1 The Side-Constraint Approach: The side-constraint approach originates with Robert Nozick (1974). On this approach, constraints take the form of *side-constraints*, which impose absolute limits on how agents may treat others. They do so by restricting the types of acts that agents may permissibly perform in the pursuit of their ends (Nozick 1974, 29). And this holds even if their end is to minimize their own performances of the restricted act-types. Indeed, side-constraints prohibit agents from infringing upon them for the sake of any end. And they function this way because of their rationale: the "Kantian principle that individuals…may not be sacrificed or used for the achieving of other ends without their consent" (Nozick 1974, 30–32).

Side-constraints serve to protect those entities that have what Kant calls *dignity*, which is a kind of value that's beyond all price (G 4:434–435).[2] Side-constraints do this by imposing strict limits on what agents may morally (and rationally) do to beings with dignity. This entails that certain act-types are never permissible, but not that they're always impermissible.[3] After all, certain morally catastrophic situations may simply be beyond morality in that morality has nothing to say about what's permissible and impermissible in such situations (Williams 1973,

[2] The 'G' stands for Kant's *Groundwork*, and the citation is given by volume and page number.

[3] Thus, 'impermissible' does not mean 'not permissible'. A chair is not permissible, but nor is it impermissible. For it's not the sort of thing that can be either permissible or impermissible.

92–93).[4] But the side-constraint view, as I understand it, entails that, in any situation in which morality is able to provide us with guidance, it is impermissible to infringe upon a side-constraint. Of course, once it's made clear that the side-constraint approach implies that there are certain types of acts that are impermissible in every non-catastrophic situation, some readers may jump to the conclusion that it's a non-starter. But this would be a mistake. Whether it's a non-starter depends on how specific the restricted act-types are. Clearly, a view that implies that all killing is prohibited is a non-starter. But the same doesn't hold for a view that implies only that very specific types of killing are prohibited. For it seems that although killing is sometimes permissible (such as when it's in self-defense), intentionally killing an innocent, non-threatening person against their will without thereby producing at least 1,000 units of impersonal goodness is never permissible.

It's important to note, then, that there are just two ways for a theory to accommodate a non-absolute prohibition against φ-ing (e.g., killing). First, the theory could hold that only certain specific types of φ-ings are absolutely prohibited. Or, second, the theory could hold that agents are required both to adopt some end that would be achieved by their not φ-ing and to weigh this against those other ends that could be achieved by their φ-ing (e.g., the end of producing more impersonal good).[5] To take the first option is to adopt the side-constraint approach; to take the second option is to adopt the teleological approach.

To better understand the side-constraint approach, it will be helpful to have a particular constraint in mind. So, let's consider the Kantian constraint against treating people as mere means, where, following Kant (G 4:428), I'll use the word 'person' as a technical term meaning 'a being with a rational nature'—that is, someone who has what Kant calls *humanity* and who is,

[4] Unfortunately, Nozick simply skirts the issue (1974, 30).

[5] Some may not like talk of ends and insist that they can accommodate a non-absolute prohibition against φ-ing simply by claiming that, when the amount of good that can be produced by φ-ing is sufficiently great, the moral reason that the agent has to produce so much goodness outweighs the moral reason that they have to refrain from φ-ing. But I don't see any substantive difference between this and the claim that, when the amount of good that can be produced by φ-ing is sufficiently great, the end of producing so much goodness is weightier than the end of refraining from φ-ing.

therefore, autonomous in the sense of being capable of employing reason to set and pursue their own ends.[6] Now, an agent treats a person as a *means* if and only if they behave toward them in a certain way for the sake of realizing some end, intending the presence or participation of some aspect of them to contribute to that end's realization.[7] And an agent treats a person as a *mere means* if and only if all the following hold: (a) they treat that person as a means, (b) that person has not given their autonomous consent—consent that's been freely given in light of all the relevant information—to being treated in this way, and (c) that person can reasonably refuse to give their autonomous consent to being treated in this way.[8] What's more, a person can reasonably refuse to give their autonomous consent to being treated in a certain way if the strongest personal reasons that they can offer against their being treated in this way are at least as strong as the strongest personal reasons that anyone else can offer for their being treated in this way.[9] Thus, if, other things being equal, I inflict a severe headache on Juan against his will to prevent several others from each being inflicted with a mild headache, I treat Juan as a mere means. For he hasn't given his consent to being treated in this way and can reasonably refuse to do so given that the strongest personal reasons that he can offer against his being treated in this way (viz., that this would result in his suffering a severe headache) are at least as strong as the strongest personal reasons that each of the others can offer for treating him in this way (viz., that this would prevent each of them from suffering a mild headache).[10] But if, other things

---

[6] As Korsgaard explains, "the distinctive feature of humanity, as such, is simply the capacity to take a rational interest in something: to decide, under the influence of reason, that something is desirable, that it is worthy of pursuit or realization, that it is to be deemed important or valuable, not because it contributes to survival or instinctual satisfaction, but as an end—for its own sake" (1996, 114).

[7] This is borrowed with modifications from Kerstein (2013, 58).

[8] Here and elsewhere in the paper, I've settled on a particular interpretation of Kantian doctrine, but the particular interpretation is not important for my overall argument.

[9] Note that I offer only one sufficient condition for reasonable refusal and so remain neutral on what other sufficient conditions there are as well as on whether any are necessary.

[10] The personal reasons that a person can offer for (or against) someone's (either themself or some other) being treated in a certain way are just the reasons they have for preferring (or dis-preferring) what their own situation would be if this someone were treated this way to what their own situation would be if this someone weren't treated this way. See de Marneffe (2013, 51).

being equal, I inflict a severe headache on Juan against his will to prevent several others from each being killed, then, although I treat Juan as *a means*, I don't treat him as *a mere means*. For he cannot, it seems, reasonably refuse to consent to his being treated in this way given how much more each of the others has at stake.

By including a constraint against treating people as mere means, a moral theory accounts for the separateness of persons. As Nozick points out, "there are only individual people, different individual people, with their own individual lives. Using one of these people for the benefit of others, uses him and benefits the others. Nothing more. …[Thus,] to use a person in this way does not sufficiently respect and take account of the fact that he is a separate person, that his is the only life he has" (Nozick 1974, 32–33). Consider, then, that when I inflict a *severe* ten-minute headache on Juan to prevent each of five others from being inflicted with a *mild* ten-minute headache, there is no conscious entity who suffers anything as bad for it as this severe ten-minute headache is for Juan. For even if a mild fifty-minute headache is just as bad as a severe ten-minute headache, there is no composite mind who suffers anything like a mild fifty-minute headache. There are only these other individual minds/persons, who each suffer one mild ten-minute headache. And, so, if we want to adequately acknowledge the separateness of persons, we must include a constraint against treating people as mere means and recognize that it is reasonable for them to refuse to consent to being so treated if what they would suffer as a result is at least as bad as what anyone else would suffer if they weren't so treated.

Although including a constraint against treating people as mere means is sufficient to account for the separateness of persons, it's insufficient to account for the Kantian idea that people have dignity in the Kantian sense. For that, we must include not merely a constraint, but a side-constraint, against treating people as mere means. For we need the no–trade-offs aspect that a side-constraints brings. We need this aspect because whereas that which has a price can be morally and rationally sacrificed, exchanged, or traded away for something of equivalent price, that which has dignity has no such equivalence and is, thus, both irreplaceable and non-substitutable. To have Kantian dignity, then, is to have what Kant calls "incomparable worth" (G 4:435–436), which is something that cannot be morally (or rationally) sacrificed, exchanged,

or traded away for anything else. Thus, only a moral theory that recognizes that persons are inviolable in that there are strict limits on what agents may morally and rationally do to them in the pursuit of their own ends (at least, in non-catastrophic situations) can account for the Kantian dignity of persons.

So, side-constraints account for the idea that persons have Kantian dignity and are, thus, inviolable. But there are different degrees of inviolability, and, so, we must ask: "To what degree are people inviolable?" The more different ways in which agents are strictly prohibited from treating people without their consent the greater their degree of inviolability. Thus, a being who may not be killed in self-defense has, other things being equal, a greater degree of inviolability than a being who may be killed in self-defense (Kamm 1996, 274). And a being who may be murdered only for the sake of saving no fewer than 1,000 lives has, other things being equal, a greater degree of inviolability than a being who may be murdered merely for the sake of saving 999 lives. (And note that I'll be using the word 'murder' as a technical term meaning 'deliberately doing something that non-consensually causes lethal harm to an innocent person when causing such harm is unnecessary to protect others from that person'.) Thus, there may not be a side-constraint against risking murder, but only a side-constraint against certain specific types of risking murder: such as those that involve, say, taking more than an n chance ($0 \leq n \leq 1$) of murdering someone for the sake of saving no fewer than $n \times 1{,}000$ lives. Thus, side-constraints should specify the precise extent to which persons are protected from various sorts of non-consensual treatment by specifying which specific act-types are prohibited. And this is what Kamm (1996, 264–75) calls a *specified* side-constraint. To violate a specified side-constraint against, say, risking murder, you must do more than simply risk committing murder; you must do so in a way that falls outside of the specified allowances—e.g., you must take more than an n chance ($0 \leq n \leq 1$) of murdering someone for the sake of saving no fewer than $n \times 1{,}000$ lives.[11]

[11] Imagine that Paola had the opportunity to spin the Wheel of Life and Death but declined to do so (see Hare 2011 for a similar case). Her spinning the wheel was guaranteed to save 499 lives but had a 0.5 objective chance of non-consensually killing Aditya, an innocent bystander who posed no threat to anyone. According to this specified side-constraint, Paola was prohibited from spinning the wheel. For given that there was a 0.5 chance that spinning it would have killed Aditya, doing so needed to ensure that at least 500 (that is, $0.5 \times 1{,}000$) lives would have been saved. And note that throughout this paper I'll be primarily concerned with what's permissible in the fact-relative

Of course, this appeal to people's limited degree of inviolability isn't the only way to account for the fact that the permissibility of risking murder depends on such things as how great that risk is, how bad it would be for that murder to take place, and how much good would come from taking that risk. For instead of holding that there is some precisely specified side-constraint that tells us exactly what kinds of risk-takings are absolutely prohibited, we might instead think that agents should adopt the end of minimizing such risks but allow that this end is sometimes to be traded off against other legitimate ends. Thus, it may be permissible to risk murder for the sake of achieving other legitimate ends. But, of course, this is no longer the side-constraint approach, but is rather the teleological approach. So, I'll turn now to explicating it.

2.2 *The Teleological Approach:* The teleological approach originates with consequentializers such as Sen (1982), Smith (2009), and Portmore (2011).[12] As they see it, act-consequentialism (hereafter, simply 'consequentialism') tells us, first, what our ends should be and, second, how we should act to suitably achieve them. Thus, on consequentialism, what our ends should be is explanatorily prior to how we should act, which is what makes it teleological. And, so, we can't easily figure out what we should do without first figuring out what our ends should be. This contrasts with theories that include side-constraints. Given that side-constraints prohibit agents from performing certain act-types for the sake of any end, we can know that we must refrain from performing such act-types without ever knowing what our ends should be.

---

sense—the sense in which it would be impermissible to do what it would in fact be bad to do even if one's evidence suggests that this is what it would be good to do. Also, throughout, I'll be primarily concerned with objective (i.e., ontic) probabilities—probabilities that are out there in the world and that would, therefore, remain the same even if we were omniscient. The objective probability that some event will (or would) occur is the percentage of the time that it will (or would) occur under identical causal circumstances—circumstances where the causal laws and histories are the same.

[12] 'Consequentialism' is a family-resemblance term that refers to all and only those theories that are, in certain important structural respects, like its archetype: classical utilitarianism (CU)—see Sinnott-Armstrong (2019) and Portmore (Forthcoming). Now, different philosophers have developed different consequentialist theories depending on what they see as CU's most attractive structural features. And they have been led to develop such theories out of a concern to avoid the counterintuitive moral verdicts that result from combining such structural features with CU's simplistic value theory: quantitative hedonism. And this is what's now known as the consequentializing project: the project of developing a moral theory that combines a consequentialist structure with a more sophisticated value theory, thereby preserving CU's attractive structural features while avoiding at least some of its counterintuitive implications.

What's more, these theories hold that what we should do is, contrary to consequentialism, explanatorily prior to what our ends should be. Thus, if there's a side-constraint against breaking a promise, we should adopt the end of never breaking a promise simply because we should never break a promise.

So, unlike theories containing side-constraints, consequentialism holds that the normative statuses of our ends are explanatorily prior to the normative statuses of our actions. But, as consequentializers have been keen to point out, consequentialists needn't hold that agents must all adopt the same set of ends. Perhaps, I should have as *my* end that I minimize the promises that *I* break, but you should have as *your* end that you minimize the promises that *you* break. So, in addition to having certain agent-neutral ends (such as minimizing promise-breakings overall), it may be that we should each also have certain agent-relative ends (such as minimizing the promises that we *ourselves* break).

Besides telling us what our ends should be, a consequentialist theory must tell us how we must act to suitably achieve them. And, of course, different consequentialists will hold different views. My own view, though, is that an act suitably achieves the ends that the agent ought to have if and only if there is no available alternative act whose prospect they ought to prefer to that of its own.[13] For it seems that which of a set prospects an agent ought to most prefer is just a function of what ends they should have and how much relative weight they should give each of them. And, so, I endorse the following form of consequentialism.

[13] An act's outcome is the way that the world would turn out if it were performed. However, if the laws of nature are indeterministic or the act itself is sufficiently vague, there needn't be *a way* that the world *would* turn out if it were performed; there may instead be only *several different ways* that it *could* turn out if it were performed. Thus, it's better to talk of an act's prospect. An act's prospect is the probability distribution consisting in the mutually exclusive and jointly exhaustive set of possible worlds that could be actualized by the given act, with each possibility assigned an objective probability such that their sum equals 1. And, of course, if there is only one possible world that could be actualized by an act, then its prospect will just be its outcome. Thus, strictly speaking, there's no need to talk about outcomes at all. We can just talk about the prospect of acts, which in certain instances (where the objective probability that some possible world will result is 1) will be the outcome of that act.

Agent-Relative Consequentialism (ARC): For any subject S and any act available to them φ, S's φ-ing is morally permissible if and only if there is no available alternative act whose prospect they ought to prefer to that of their φ-ing. And the most fundamental permissibility-making feature of a permissible action is its lacking an available alternative whose prospect they ought to prefer to that of its own.[14]

ARC allows us to accommodate constraints within a teleological framework. To illustrate, take the constraint against breaking a promise. To accommodate this constraint, ARC need only hold that, for any subject S, S ought to prefer the prospect of their refraining from breaking a promise to the prospect of their breaking that promise to prevent two others from each breaking a morally comparable promise. The idea is that subjects ought, other things being equal, to give greater weight to the end of minimizing their own promise-breakings than to the end of minimizing promise-breakings overall.[15] Thus, the resulting version of ARC would imply, for instance, that Abbey is prohibited from breaking her promise to Abe even to prevent both Bertha from breaking her morally comparable promise to Bert and Carla from breaking her morally comparable promise to Carl. The resulting view would also imply that Carla is prohibited from breaking her promise to Carl even to prevent both Abbey from breaking her morally comparable promise to Abe and Bertha from breaking her morally comparable promise to Bert. Indeed, it yields the same constraint against breaking a promise regardless of who's in

[14] Strictly speaking, what I endorse is a maximalist, dual-ranking revision of ARC. See Portmore (2011; 2019). But these qualifications won't matter here.

[15] On an alternative teleological view, we would accommodate constraints by adopting agent-neutral consequentialism and holding that agents are required both to have the end of minimizing promise-breakings-committed-in-order-to-prevent-other-promise-breakings and to give this end priority over that of minimizing promise-breakings overall—see Setiya (2018, 97–99) and Dougherty (2013, 531). I, like some others (e.g., Howard 2021, 741), find this claim about our obligatory ends implausible. Indeed, it seems especially implausible given that, as even Setiya (2018, 96) suggests, each of us ought to prefer that it is, other things being equal, someone else rather than ourself who fails to keep a promise. And see Cox and Hammerton (Forthcoming) for a related worry. What's more, as we'll see in section 4 below, the sort of view that Setiya and Dougherty propose yields counterintuitive moral verdicts in what are known as *intra-agent minimizing cases*. For all these reasons, I'll ignore this alternative in the remainder.

the position of breaking a promise to minimize morally comparable promise-breakings.

What makes this approach to incorporating constraints teleological is that it treats not performing an act of a certain type as an end and, so, it allows—at least, in principle—that the agent's other legitimate ends can affect the permissibility of her performing this act-type.[16] Thus, although this view prohibits breaking a promise to prevent two others from each breaking a morally comparable promise, it can permit doing so for the sake of achieving various other ends. And, so, it can accommodate the following intuitive moral verdicts: (V1) it's permissible to break a promise to help a friend move to prevent someone from losing a limb; (V2) it's permissible to break a promise to help a friend move to prevent a million others from each breaking a morally comparable promise; (V3) it's permissible to break a promise to help a friend move on this present occasion to prevent oneself from breaking a morally comparable promise on each of five future occasions; (V4) it's permissible to break a promise to help a friend move to reduce by half the risk of someone's being murdered; and (V5) it's permissible to break a promise to help a friend move to reduce by one-tenth the risk that one will commit murder. After all, it seems that agents should have several other ends besides that of ensuring that they themselves do not break, say, the promise to help a friend move on this present occasion, including all the following: (E1) minimizing promise-breakings overall; (E2) minimizing their own promise-breakings over time; (E3) minimizing the risk to each person that they will lose a limb; (E4) minimizing the risk to each person that they will be murdered; and (E5) minimizing the risk that they themselves will commit murder. What's more, it seems that an agent should prefer the prospect of their doing what's necessary to achieve these ends to that of their keeping their present promise to help a friend move, for it seems that ends such as E1–E5 should be given more weight than that of keeping such a promise. Thus, ARC can accommodate a wide of

[16] Of course, the teleological approach could yield the same moral verdicts that the side-constraint approach does by holding that the end of not performing an act of a certain type (say, murder) is to be given lexical priority over all other ends. In that case, the teleologist would claim that it is never permissible to commit murder just as those who claim that there is a side-constraint against murder do. But the two views would still differ in their accounts of why it's never permissible to commit murder. For the teleologist would hold that it's because this end has lexical priority over all other ends, whereas the side-constraint theorist would hold that it's because there are certain types of acts that we're prohibited from performing in the pursuit of any end.

range of intuitive moral verdicts concerning when it is and isn't permissible to break a given promise.

3. Three Advantages to Taking the Teleological Approach

3.1 Its Theoretical Simplicity. Besides being able to accommodate a wide range of intuitive moral verdicts, the teleological approach has the advantage of being able to do so with less theoretical apparatus. For even those who reject the teleological approach should accept two of its claims: (C1) that agents should have ends such as E1–E5 and (C2) that they should, other things being equal, do what would best achieve them. Thus, what sets plausible versions of the side-constraint approach apart from the teleological approach is not that it denies these two claims, but only that it makes the additional claim that there are side-constraints that limit what agents may permissibly do in pursuit of these ends. The problem, though, is that once we accept these two claims, there's no need to incur this additional commitment. For, as I'll show below, these two plausible claims are themselves sufficient to account for all the relevant intuitive moral verdicts.[17]

The side-constraint theorist should accept C1—that agents should have ends such as E1–E5—to account for why agents should feel some lingering regret whenever they're forced to sacrifice one of these ends. To illustrate, consider that agents are morally required to break a relatively trivial promise to save a life. Yet, someone who did so should still regret having to break their promise. For they should want not only to save lives but also to keep their promises. And this end is what explains—at least, in part—why they should feel some residual dissatisfaction in having to break their promise to save the life even though they are permitted to do so.

Also, the side-constraint theorist should accept C2—that agents should, other things being equal, do what would best achieve the ends that they ought to have—to explain why

[17] But, perhaps, the thought is that, although this additional commitment isn't needed to yield any intuitive verdicts, it is needed to provide the correct explanation for them. I'll debunk this thought in section 4 below.

agents should further such ends even when doing so is unnecessary to avoid infringing upon any constraints. To illustrate, suppose that an agent must choose between doing nothing and saving a life, and assume that everything else is equal. Thus, assume that neither option would involve infringing upon a constraint. To explain why they ought to save the life, the side-constraint theorist must appeal both to the fact that agents should have saving lives as an end and to the fact that they should, other things being equal, do what would best achieve the ends that they ought to have.

Thus, it seems that the side-constraint theorist should accept both C1 and C2. Only then can they plausibly account both for the regret that agents should feel whenever they're required to sacrifice certain ends and for the *pro tanto* moral obligation that they have to achieve these ends. But, in that case, the side-constraint theorist gains nothing in terms of the verdicts that they're able to accommodate by making the additional claim that there are side-constraints. For, as the consequentializers have shown, we can accommodate all the relevant intuitive moral verdicts simply by postulating that agents should both adopt certain ends and do what would best achieve them. To illustrate, suppose that the relevant intuitive moral verdicts include V1–V5 from above. To accommodate such verdicts, the side-constraint theorist must postulate some elaborately specified side-constraint against breaking a promise, one that incorporates all the necessary allowances for accommodating such an array of verdicts. By contrast, the teleologist need only postulate what the side-constraint theorist should already accept: various obligatory ends (such as E1–E5) and a *pro tanto* moral obligation to do what would best achieve them.

3.2 Its Conception of Choice-Worthy Action. Another advantage of the teleological approach is that it, unlike the side-constraint approach, embodies a certain attractive conception of choice-worthy action. This conception stems from the realization that it is through our actions that we attempt to affect the way the world goes. Whenever we face a choice of what to do, we also face a choice of which of various possible worlds to attempt to actualize. Moreover, whenever we act intentionally, we act with the aim of making the world go a certain way. The aim needn't be anything having to do with the causal consequences of the act. The aim could be nothing more than to perform the act in question. For instance, one can run merely with the aim of running.

The fact remains, though, that for every intentional action there is some end (or ends) at which the agent aims.[18] It's natural, then, to think that the most choice-worthy act—the act that the agent ought to perform—is just the one that will make the world go as she ought to aim for it to go.[19]

This natural thought is somewhat similar to what Scheffler (1985) calls the *maximizing conception of rationality*. According to this conception, "if one accepts the desirability of a certain goal being achieved, and if one has a choice between two options, one of which is certain to accomplish the goal better than the other, then it is, *ceteris paribus*, rational to choose the former over the latter" (1985, 414). However, the thought that I have in mind is a bit different and, I believe, more plausible. For one, Scheffler's conception is beholden to what the agent takes to be the case rather than what is in fact the case, for it concerns subjective rationality rather than objective choice-worthiness. For another, Scheffler's conception appeals to the goodness (or, as he puts it, the 'desirability') of a goal's being achieved rather than the degree to which the agent should want it to be achieved. This is important, because, as Philippa Foot notes, it can "be right to prefer a worse state of affairs to a better" (1985, 198). For although we should certainly want the world to change for the better, this isn't all that we should want. We should, it seems, also want those to whom we have close ties to fare well and not solely to the extent that their faring well would promote the impersonal good. Thus, I should prefer that my daughter rather than some stranger is saved even if it would be slightly better, impersonally speaking, were the stranger saved instead. It seems, then, that agents should sometimes prefer a worse state of affairs to a better. And, given this, we should think that what agents should do tracks what they should most desire rather than what's most valuable/desirable.

---

[18] This is widely endorsed by consequentialists and non-consequentialists alike. For instance, John Stuart Mill says, "all action is for the sake of some end" (*Utilitarianism*, chap. 1). And Immanuel Kant says, "an end is an object of the free faculty of choice [i.e., the free Willkür], the representation of which determines it to action, whereby [the end] is brought about. Every action, thus, has its end" (*Metaphysics of Morals,* 6:384–385).

[19] Even some non-consequentialists come close to endorsing this. For instance, Judith Jarvis Thomson says: "given that for a person to act just is for the world to go in a way that it otherwise would not go, surely the question whether he ought to act had better turn on a comparison between how it will go if he acts and how it will go if he does something else—to repeat, there seems to be nothing else for it to turn on" (2003, 8).

For these reasons, I believe that we should accept, not what Scheffler calls the *maximizing conception of rationality,* but what I'll call the *maximizing conception of choice-worthy action*: if one ought to have a certain set of ends, and if one has a choice between two options, one of which will better achieve these ends than the other, then one ought to choose this over the other.[20] This, I believe, is quite difficult to deny.[21] Consequently, it's hard to see how Abbey ought to refrain from breaking her promise to Abe to prevent both Bertha from breaking her morally comparable promise to Bert and Carla from breaking her morally comparable promise to Carl if the only end that Abbey should have is to minimize promise-breakings overall. And, thus, such a constraint can seem paradoxical unless we admit that Abbey should additionally have some end that would be achieved by her refraining from breaking her promise to Abe, such as either the end of refraining from breaking any promises at present or the end of minimizing her promise-breakings over time. For if actions aim at achieving ends, then although it would make sense to prohibit performing certain types of actions when performing such actions would thwart these ends, it wouldn't make sense to prohibit such actions even when performing them would best achieve the ends that one ought to have. But we can avoid such paradoxical views by embracing the maximizing conception of choice-worthy action and using the teleological approach to accommodate constraints on ARC.

3.3 Its Ability to Plausibly Deal with Morally Relevant Indeterminacy. Lastly, ARC is better suited to deal with the fact that there can be indeterminacy with respect to whether an agent has infringed (or would infringe) upon a given constraint. But, before explaining what makes ARC better suited to deal with this, let me explain how such indeterminacy can arise. There are at least three ways.

---

[20] This is because the reason to perform an action is just that it would further the ends that one has reason to pursue. As Allen Wood, interpreting Kant, puts it: "an action is something that is in the agent's power, and is chosen as a means to the agent's end (G 4:427). Looked at from this standpoint, it is the end that supplies the reason for every action. But…there must be a reason for setting the end. The end [must be] good or valuable in some way; …or it [must be] an end that morality requires you to have…. So instead of saying only that the end is the reason for the action, it is more appropriate to say that the reason for setting the end is the reason for the action" (2017, 266).

[21] Admittedly, however, some do. See Hurley (2018) and Muñoz (2021).

First, whether an act counts as infringing upon a constraint can depend on what some free agent would have done had that act not been performed. To illustrate, assume that there's a constraint against murder. (And recall that I'm using the word 'murder' to mean 'deliberately doing something that non-consensually causes lethal harm to an innocent person—that is, a person who has not yet committed any offense—when causing such harm is unnecessary to protect others from that person'.) And now consider *The Sniper*. Imagine that Gunnar—an innocent person—was shot and killed by a sniper as he was attempting to enter a school building with a gun. Did the sniper infringe upon the constraint against murder in killing Gunnar? Well, it depends on whether this was necessary to protect others, which in turn depends on whether Gunnar would have hurt anyone had he not been shot. But it's entirely possible that he had the kind of libertarian freedom that implies that there was some ontic probability between zero and one that he would have hurt someone had he not been shot. And, in that case, there will be no fact of the matter as to what he would have done had he not been shot. And, so, it will be indeterminate whether the sniper infringed upon the constraint against murder.

Second, it could be indeterminate whether some necessary condition for infringing upon a constraint has been satisfied. To illustrate, assume that there's a constraint against breaking a promise. And now consider *The Promise*, which is a revised version of a case given by Jackson and Smith (2006). Imagine that there is just no fact of the matter as to whether a woman named Lupe promised to vote for Candidate X at last night's party. Perhaps, she said "I promise to vote for Candidate X" but it's ontically indeterminate whether she was too drunk at the time to be capable of making a genuine promise—after all, there will certainly be a sorites series from a BAC (blood alcohol content) of 0.0 to a BAC of 0.37, which is potentially lethal. Or, perhaps, she said something like "I will vote for Candidate X" but the context left it ontically indeterminate whether she was thereby promising to vote for Candidate X or just predicting that she would vote for Candidate X. Nevertheless, suppose that she ends up refraining from voting for Candidate X. In so refraining, did she break a promise? It seems indeterminate. For it's ontically indeterminate whether she even promised to vote for Candidate X. And, thus, it's ontically indeterminate whether she has infringed upon the constraint against breaking a promise.

Third, indeterminacy with respect to whether an agent would have infringed upon a constraint had they performed a certain act can arise given that that act counts as infringing upon the constraint if and only if a certain counterfactual is true and counts as not infringing upon that constraint if and only if its contradictory counterfactual is true. For, in some instances, neither counterfactual will be true given the semantics of counterfactuals.

To illustrate, assume that there's a constraint against murder (as stipulatively defined above). And, now, consider the following case, which I borrow, with some revisions, from Jean-Paul Vessel (2003, 104–5).

> *The Demon's Coin:* A powerful demon produces a magical but fair coin and offers Parisa the chance to flip it. If she flips the coin and it lands heads, that will cause lives to be saved. If she flips the coin and it lands tails, that will cause lives to be lost. And, if she abstains from flipping it, things will remain the same. Parisa decides to abstain.

The relevant counterfactuals are:

CF1 If Parisa had flipped the coin, it would have landed heads.

CF2 If Parisa had flipped the coin, it would have landed tails.

If CF1 is true, then Parisa would not have violated the constraint had she flipped the coin. And, if CF2 is true, then she would have violated the constraint had she flipped the coin. But which is true? Clearly, they can't both be true. Since they have the same antecedent and logically contradictory consequents, one will be false if the other is true. (Assume that the coin must land heads or tails.) Interestingly, though, neither is true. They're either both false or both indeterminate.[22] This is because there is no fact of the matter as to whether the coin would have

[22] On Lewis's theory (1973), they're both false. And, on Stalnaker's theory (1984), they both have indeterminate truth values. And these two theories seem to be the two leading contenders. (I'm told, though, by Daniel Muñoz in

landed heads (or tails) had Parisa flipped it. To accept this, we don't need to assume the laws of nature are indeterministic. Nor do we need to hold that there is more than just one actual future. We need only to accept both that (1) the antecedents in CF1 and CF2 are underspecified given that there are countless specific ways that Parisa could have flipped the coin (only some of which would have resulted in the coin's landing heads) and that (2) Parisa lacks the ability to determine whether she flips the coin in any of the specific ways that would result in its landing heads.

Parisa lacks the ability to determine whether she flips the coin in any of the specific ways that would result in its landing heads, because whether it lands heads depends on very minute differences with respect to how she flips it: e.g., the precise locations of both the coin and her thumb when the two make impact, the precise force with which her thumb impacts the coin, and the precise orientation of both the coin and her thumb on impact. Since she lacks the dexterity to determine these details with any precision and since the result of her coin toss is extremely sensitive to them, she can't control whether she flips the coin in a way that results in its landing heads. And, given this, there is just no fact of the matter concerning whether, had she flipped the coin, it would have landed heads. And, thus, there is no fact of the matter concerning whether, had she flipped the coin, she would have infringed upon the constraint against murder.

These three are cases of *morally relevant indeterminacy*—that is, indeterminacy with respect to whether some morally relevant state of affairs obtains (or would obtain). Given all these different ways that moral indeterminacy can arise, I doubt that we can avoid it altogether. Of course, some might reply that, in *The Promise*, the relevant indeterminacy is only semantic and not ontic, claiming that although we can't know whether Parisa was too drunk to be capable of making a genuine promise, there was, nevertheless, some fact of the matter (see Williamson 1994). And others might reply that, in *The Sniper*, the relevant constraint can't be against murder but must instead be against taking an objective risk of committing murder.

---

correspondence that Alan Hájek is developing a new theory, but, like Lewis's theory, it is one on which CF1 and CF2 are both false.)

Now, I'm not convinced that such replies succeed. But, even if I were to concede that they do, there would still be cases like *The Demon's Coin*, where the morally relevant indeterminacy seems unavoidable. Consequently, I think that our moral theories must be prepared to deal with such morally relevant indeterminacy.

One way of dealing with morally relevant indeterminacy is to just allow that wherever there is morally relevant indeterminacy, there must also be *deontic indeterminacy*—that is, indeterminacy regarding an act's deontic status (such as whether it ought to be performed).[23] But there are at least two reasons to reject any moral theory that countenances deontic indeterminacy (and both are inspired by similar remarks in Dougherty 2016).

First, prospectively, moral theories are meant to tell us what we ought to do. But if there is deontic indeterminacy, then it will be as if we are presented with a tile that is neither red nor not red and commanded to put it in a certain box if and only if it is red and to refrain from putting it in that box if and only if it is not red (Dougherty 2016, 449). The problem is that there's no way for us to do as commanded, which makes such commandments pointless. Imagine, then, that you have the option of φ-ing and a given moral theory implies that it's indeterminate whether you morally ought to φ. What's more, assume that this theory—as any moral theory must—directs you to φ if and only if φ is what you morally ought to do, and directs you to refrain from φ-ing if and only if φ is not what you morally ought to do. The problem is there is no way for you to do as directed. And what's the point of a moral theory that gives you (and the rest of us) directives that are impossible to follow?

---

[23] Several philosophers have argued that deontic indeterminacy just follows from morally relevant indeterminacy—see, for instance, Dougherty (2016, 449), Schoenfield (2016, 262-3), and Dunaway (2017, 40). They all give the following sort of argument based on a sorites series: "I will assume that it is determinately permissible to terminate a one-day old zygote [and determinately impermissible to terminate a one-year old child]. However, there seems no specific point in an entity's continuous development from zygote to one-year old at which it acquires moral personhood. …Instead, the entity passes through a range of borderline cases of moral personhood. When the entity is in this range, it is indeterminate whether it is permissible to terminate it" (Dougherty 2016, 449). But from the fact that it's permissible to terminate a non-person, impermissible to terminate a person, and indeterminate whether a certain entity is a person, it doesn't follow that it's indeterminate whether it's permissible to terminate that entity. For a moral theory could hold that it's not only impermissible to terminate a person, but also impermissible to terminate any borderline case of moral personhood.

Second, retrospectively, moral theories are meant to tell us whether morally responsible agents are blameworthy (or praiseworthy)—and, thus, deserving of sanction (or reward)—for their actions. But if a moral theory allows for deontic indeterminacy, then it will sometimes be indeterminate whether a morally responsible agent has done anything blameworthy (or praiseworthy). For instance, if, in *The Demon's Coin*, it's indeterminate whether Parisa acted wrongly in refraining from flipping the coin (because it's indeterminate whether her flipping the coin would have saved or cost lives), then it will be indeterminate whether she is blameworthy or praiseworthy in so refraining. And, thus, it will be indeterminate whether she deserves sanction or reward. But we might wonder what's the point of a moral theory that can't tell us whether morally responsible agents who perform non-neutral acts are blameworthy or praiseworthy for their behavior. What's more, we might wonder whether it even makes sense to suppose that it could be indeterminate whether someone deserves sanction or reward for responsibly performing a non-neutral act, for this doesn't seem to be the sort of thing on which an adequate moral theory can just remain silent.

It's fortunate, then, that ARC can deal with morally relevant indeterminacy without having to countenance deontic indeterminacy. To illustrate, consider again *The Promise*. Although there is no fact of the matter as to whether Lupe broke a promise in refraining from voting for Candidate X, ARC can insist that in refraining from voting for Candidate X she acted either permissibly or impermissibly, and determinately so. For ARC can insist that Lupe is required to have as one of her ends that it be (determinately) true that she has not broken a promise. And she achieves this end if and only if she votes for Candidate X. Of course, she ought to have other ends as well, such as that of making the world better. And let's suppose that she did make the world better in refraining from voting for Candidate X given that this resulted in a better candidate's being elected. And, so, whether Lupe acted permissibly in refraining from voting for Candidate X just depends on the relative importance of these two ends: (1) making the world go better by refraining from voting for Candidate X and (2) making it (determinately) true that she hasn't broken a promise by voting for Candidate X. If the former is at least as important as the latter, then she acted permissibly in refraining from voting for Candidate X and, if not, she acted impermissibly. Thus, ARC will hold that even though it's

indeterminate whether Lupe promised to vote for Candidate X, it's determinate whether she acted permissibly in refraining from doing so. So, the teleological approach can handle morally relevant indeterminacy without having to countenance deontic indeterminacy.[24]

By contrast, the side-constraint approach must countenance deontic indeterminacy if there are side-constraints against such things as murder and promise-breaking. For, as we've seen above, it can be indeterminate whether an act is of such a type. Of course, the side-constraint theorist may just deny that there are side-constraints against performing such act-types and hold instead that there are only constraints against doing what will make it (determinately) true that one has performed an act of this type. That is, they could deny that there is, say, a constraint against breaking a promise and hold instead that there is only a constraint against doing what would make it (determinately) true that one has broken a promise. This is what we might call the *truth-centric* side-constraint approach (Williams 2017). This approach may seem odd and perhaps even *ad hoc*, but it would at least allow the side-constraint theorist to avoid deontic indeterminacy. Nevertheless, I find it implausible. To see why, consider again *The Promise*. In this case, the fact that Lupe did and said things last night that make it indeterminate whether she promised to vote for Candidate X seems like a good moral reason for her to vote for Candidate X. But it's hard for the truth-centric side-constraint approach to account for this. After all, there is, on this approach, only a constraint against doing what will make it (determinately) true that she has broken a promise. Thus, there is on this view no constraint that Lupe infringes upon by refraining from voting for Candidate X. Why, then, is there a moral reason for her instead to vote for Candidate X? The answer, it seems, is that Lupe ought both to have ensuring that it's (determinately) true that she has broken no promises as an

[24] The reader may wonder about higher-order indeterminacy. For instance, the reader may wonder what the teleologist would say if it's indeterminate whether in refraining from voting for Candidate X it's indeterminate that she has broken a promise. Perhaps, there's a 0.4 objective chance that it's indeterminate and a 0.6 objective chance that it isn't. But, even in such a case, there can still be a determinate fact of the matter concerning whether Lupe ought to prefer the prospect of her voting for Candidate X to the prospect of her refraining from voting for Candidate X given her obligatory and discretionary ends and the objective probabilities that they will be achieved on each of these two alternatives. And, consequently, there will, on the teleological approach, be a determinate fact about whether she's permitted to refrain from voting for Candidate X even if it's indeterminate whether her so refraining would make it indeterminate that she has broken a promise.

end and to do what will, other things being equal, best achieve the ends that she ought to have. In other words, we see again that the side-constraint theorist should accept claims such as C1 and C2. But, as we saw above, they do so at some cost. For once they do so, there's no longer any reason to postulate side-constraints, for these claims are themselves sufficient to account for all the intuitive verdicts. Thus, the side-constraint theorist can avoid deontic indeterminacy only at the cost of theoretical simplicity.

4. Two Putative Disadvantages to Taking the Teleological Approach

I've argued that the teleological approach has several advantages over the side-constraint approach. Of course, this doesn't mean that the teleological approach wins the day. For it may be that the teleological approach has its own disadvantages, and these could tip the balance back in favor of the side-constraint approach. Indeed, some have thought so (see, e.g., Brook 1991, Emet 2010, Howard 2021, Kamm 1996, Löschke 2020, and Otsuka 2011). But, as I'll argue below, their thinking is the result of misunderstanding what the teleologist is committed to.[25]

4.1 Some claim that the teleological approach fails to acknowledge both the true value of persons and the fact that some constraints are ultimately grounded in the statuses of patients as opposed to the ends of agents: First, some philosophers (e.g., Anderson 1993) have objected to consequentialist views such as ARC on the grounds that they fail to acknowledge the true value of persons. As they see it, persons are among the fundamental bearers of value. Indeed, they have inherent, unconditional, and irreplaceable value in virtue of their distinctive capacity for employing reason to set and pursue their own ends. And we respond to this value appropriately by first and foremost respecting them and their rational choices. For this respect is owed to them. But, as these philosophers see things, consequentialists must deny this. As they see it, consequentialists are committed to there being only one kind of value: the kind that's to be

[25] One putative disadvantage of the teleological approach is that it is "gimmicky" (Nozick 1974, 29). I won't address this worry here, for it has been adequately dealt with elsewhere—see, e.g., Broome (1991), Dreier (1993), Sen (1983), Smith (2009), Otsuka (2011), and Vallentyne (1988). What's more, I believe that the version of consequentialism that I develop below will seem anything but gimmicky.

desired and, consequently, promoted. And since only states of affairs can be promoted, they see consequentialists as being committed to states of affairs being the only fundamental bearers of value. Thus, they see consequentialists as committed to people having only conditional value: value solely on the condition that they make the realization of good states of affairs possible. They assume, then, that consequentialists must regard people as mere receptacles for value (e.g., pleasure), which makes them replaceable in that we may permissibly destroy some of them for the sake of bringing others into existence so long as there will, in the end, be at least as much good held in these new receptacles as were held in the old ones.

Second, several philosophers (e.g., Emet 2010, Howard 2021, and Löschke 2020) have objected to such consequentialist views on the grounds that they fail to acknowledge the fact that some constraints are ultimately grounded in the statuses of patients as autonomous beings whose rational choices are owed respect. For, as they see it, consequentialists must instead hold that the ultimate grounds for constraints lie with agents and their self-centered desire to keep their own hands clean. For instance, Howard claims that the consequentialist's defense of the wrongness of your infringing upon a constraint is "self-indulgent," for "it appeals ultimately to a self-centered preference to keep your own hands clean" (2021, 246). By contrast, Howard and others believe that constraints are "only plausibly defended by referencing not some fact about you, the agent, but about the patient. [For, as they see it,] it is in virtue of something about the patient, …the respect you owe her…, that generates the agent-relative reasons against" your infringing upon the constraint (Emet 2010, 6). And, thus, as they see it, constraints are "victim-focused rather than agent-focused" (Kamm 1996, 279) in that they are grounded only in the statuses of the potential victims and not in the ends of the agents.

But both objections are simply the result of these philosophers misunderstanding what consequentialists are committed to. Consequentialists are committed to the evaluative ranking of prospects being explanatorily prior to the deontic statuses of the actions associated with them, but this doesn't commit them to valuing persons in any particular way.[26] Nor does it

[26] Thus, if you wish to eschew consequentialism altogether while justifying constraints against treating people in certain ways by appeal to their humanity, then you should understand the normative significance of their humanity in non-axiological terms—see Bader (Forthcoming). (And note that an evaluative ranking of some set is one that

commit them to holding that it is the ends of agents as opposed to the statuses of patients that *ultimately* ground constraints. For although the consequentialist is certainly committed to the deontic statuses of actions being grounded in the evaluative ranking of their prospects, they needn't deny that this evaluative ranking is ultimately grounded in the statuses of patients.[27] To demonstrate this, I will now develop a Kantian version of ARC, one that accepts as much as possible a Kantian theory of the value of persons. And I'll call it *Kantsequentialism*.[28]

Kantsequentialism endorses the following Kantian theory of value. The fundamental bearers of value include persons. And persons are ends-in-themselves—that is, beings who have objective worth in virtue of their inherent nature. Importantly, the value of an end-in-itself is most immediately normative for non-propositional attitudes such as love and respect rather than for propositional attitudes such as 'desiring that' and 'intending to bring it about that' (Anderson 1993, 17). Indeed, we should desire and pursue certain states of affairs mainly as a way of properly expressing our respect for people and their rational choices (Anderson 1993, 20). What's more, people are, in Kant's terminology, *self-standing ends*. "A self-[standing] end (selbständiger Zweck) (G 4:437)…is something that already exists and whose 'existence is in

---

ranks its members in terms of the weight of the subject's reasons for desiring each of them—see Portmore 2011, 34–38.)

[27] Of course, you may define 'consequentialism' differently. As you define it, consequentialism may necessarily be an agent-neutral theory (Howard-Snyder 1993). Or it may, on your definition, be committed to holding that states of affairs are the only fundamental bearers of intrinsic value (Anderson 1993, 30). Or it may even be that, on your definition, consequentialism cannot give a foundational role to anything other than the evaluative ranking of prospects—and, thus, cannot, as Kantsequentialism does, give a foundational role to our duty to respect persons. Fair enough. 'Consequentialism' is after all a family-resemblance term (see Sinnott-Armstrong 2019 and Portmore Forthcoming), and people will, therefore, define it differently depending on what they take the most important feature of its archetype—viz., classical utilitarianism—to be. Now, as I see it, the most important feature of classical utilitarianism is that it takes the deontic statuses of actions to be grounded in an evaluative ranking of their prospects. But, in any case, how we label theories such as ARC is not that important. Indeed, what's important for my purposes is only that ARC can adopt the teleological approach to accommodating constraints and thereby gain its advantages over the side-constraint approach.

[28] I borrow this clever term from Richard Arneson's 1997 PHIL 224 class handout entitled "Consequentialism and Justice." Note that Kantsequentialism isn't at all like Cummiskey's Kantian consequentialism. For whereas Cummiskey (1996) attempts to derive agent-neutral consequentialism as a first-order moral theory from Kant's metaethical assumptions, I will attempt to incorporate Kantian first-order moral verdicts within a consequentialist framework.

itself an end', having worth as something to be esteemed, preserved, and furthered" (Wood 1999, 115). And a self-standing end contrasts with what Kant calls "*an end to be effected* (zu bewirkender Zweck), that is, some thing or state of affairs that does not yet exist but is to be brought about through an agent's causality (G 4:437) [italics added]" (Ibid). Thus, people have the sort of value that's to be respected when present rather than brought about when absent. And, so, this value doesn't call for us to bring as much of it as possible into existence but, rather, to respect what already has it. Indeed, we *owe* those who have it our respect. What's more, the value of persons is unconditional. And, thus, a person cannot lose their worth as an end-in-themself except by ceasing to be a person (that is, by ceasing to have the rational capacity for setting and pursuing their own ends). So, even if a person commits horrible deeds or loses their capacity for happiness, their worth as an end-in-themself is not at all diminished.

Again, the proper evaluative response to such value is first and foremost to respect that which has it. And this is an attitude, not an action. But, like all attitudes, this attitude is associated with certain motivational tendencies and the actions that they commonly engender. Thus, respect for persons will involve not only having the affect associated with feeling awe and esteem for their humanity, but also having the following dispositions: (1) not to interfere with their autonomous choices, (2) to inquire with an open mind as to the reasons for their choices, (3) to try to reason with them rather than manipulate them when we fear that they will otherwise make bad choices, and (4) to hold them accountable when they do make bad choices. Most importantly, though, respect for persons will involve having a disposition to refrain from treating them as mere means.

So, on Kantsequentialism, people are ends-in-themselves who are inherently and unconditionally valuable and are, thus, owed respect. Also, given that they are self-standing ends, they are irreplaceable. Consequently, it would be impermissible to kill one person even for the sake of bringing several other happier people into existence. Moreover, on Kantsequentialism, our most fundamental duty is to respect people and their statuses as ends-in-themselves. And, from this duty, we derive a duty to adopt as an end treating them always as ends-in-themselves and never as mere means. And, on Kantsequentialism, this end will have

some sort of general but non-lexical priority over both our discretionary ends and our other obligatory ends, including those of both minimizing murders overall and minimizing the murders that we ourselves commit.[29] Thus, agents ought generally to prefer the prospect of their refraining from murdering someone at present to the prospect of their doing so as a means to minimizing murders overall, or even as a means to minimizing the murders that they themself commit over time. But, because this priority is only some sort of general priority, Kantsequentialism needn't prohibit you from treating someone harmlessly as a mere means when this is necessary to prevent yourself from very likely accidentally killing someone in the future. And, because this priority is non-lexical, the quantities matter. So, if murdering someone at present is the only way to prevent, say, hundreds of murders by others or, say, dozens of future murders by yourself, then you ought to prefer the prospect of your committing that present murder to that of your refraining from doing so. Consequently, there will, on Kantsequentialism, be a non-absolute prohibition against murder such that you are permitted to murder one person as a means to preventing either more than m murders by others or more than n murders by yourself ($m > n$), but are prohibited from doing so as a means to preventing either m or fewer murders by others or n or fewer murders by yourself.

All this is compatible with acknowledging that people are ends-in-themselves who are inherently and unconditionally valuable. Indeed, the only part of Kant's theory of value that the Kantsequentialist must reject is his claim that people have "dignity" in the sense of having incomparable worth. Of course, the Kantsequentialist will accept that people have dignity in the ordinary sense of being worthy of respect, but the Kantsequentialist must deny that people have incomparable worth such that they cannot be morally (or rationally) sacrificed, exchanged, or traded away for anything else. Instead, the Kantsequentialist will endorse the teleological approach, where refraining at present from treating someone as a mere means is an end that can be sacrificed, exchanged, or traded away for the sake of sufficiently weighty other ends. But this

[29] As noted above, we could on the teleological approach—and, thus, on Kantsequentialism—give some ends lexical priority over all others and thereby mimic the moral verdicts that the side-constraint approach yields. But I just don't find this plausible. Thus, as I conceive of it, Kantsequentialism holds that various ends have only non-lexical priority over others.

is not a problem for Kantsequentialism, for Kant's claim that persons have incomparable worth is quite implausible. After all, if people had incomparable worth, then killing in self-defense would be impermissible.[30] For, in doing so, you would be killing them as a mere means to preserving your own life, sacrificing their humanity for the sake of preserving your own. But it's absurd to think that killing in self-defense is impermissible. So, we should reject the idea that people have incomparable worth such that it is never permissible to treat them as mere means. And, therefore, I believe that Kantsequentialism can acknowledge the true value of persons, for their true value is inherent and unconditional, but not incomparable.

The Kantsequentialist can also acknowledge that some constraints are ultimately grounded in the statuses of patients rather than the ends of agents. Howard (2021) denies this. For when speaking about the reasons agents have to refrain from infringing upon people's rational wills, he claims that "since these reasons to act have their source in the value of particular people, ...they're reasons to act which don't derive from...reasons for preferring some outcomes to others" (2021, 749). But this reasoning is fallacious. From the fact that A ultimately derives from (and, thus, has its source in) C, it doesn't follow that A doesn't derive from B. For it could be that A derives from B, which in turn derives from C. Indeed, that's how it is on Kantsequentialism. The ultimate source of the constraint against treating people as mere means is the value that people have. And, from this value, the Kantsequentialist derives a duty to respect them and their rational choices. In turn, the Kantsequentialist derives from this duty to respect people a duty to adopt the end of not treating them as mere means.[31] And, since Kantsequentialism gives this end general and non-lexical priority over other ends, we get a constraint against treating people as mere means. But the ultimate source of this constraint is

[30] See Kerstein (2013) for other cases that make trouble for Kant's claim that people have incomparable worth.

[31] From the fact that people have a certain kind of goodness, it follows that it's right to respect them. From the fact that it's right to respect them, it follows that it's right to adopt not treating them as mere means as an end. And, from the fact that it's right to adopt not treating them as mere means as an end, it follows that, other things being equal, the prospect of doing something that risks treating someone as a mere means ranks lower, evaluative speaking, than the prospect of doing something that doesn't have this risk. So, does the right have priority over the good, or vice versa? There's just no straightforward answer just as there wouldn't be if acts were right because they produced good outcomes and outcomes were good because it was right to desire them.

the value and status of persons as ends-in-themselves. It's just that Kantsequentialism both denies that this status is inviolable and holds that the derivation of the constraint against treating people as mere means from their statuses as ends-in-themselves proceeds via an intermediary duty to perform the act whose prospect you ought to most prefer.[32]

Of course, Howard would object to this intermediary duty on the grounds that it's "self-indulgent" in that "it appeals ultimately to a self-centered preference to keep your own hands clean" (2021, 246). But this too is a mistake. In accounting for the constraint against harming people as mere means, Kantsequentialism does not appeal to a preference to keep your own hands clean (or even as clean as possible), but rather only to a preference for not treating people as mere means—even as a mere means to minimizing the instances in which you've harmed people as mere means. To illustrate, consider *The Antidote*.[33] You and Yasmin each have a vial with two cubic centimeters (cm^3) of slow-acting poison. Although 1 cm^3 is enough to kill a person, you give Valentino all 2 cm^3. Yasmin, by contrast, gives 1 cm^3 to Vladimir and 1 cm^3 to Vasily. Shortly afterwards, you are given a vial containing 2 cm^3 of the only available antidote. It takes an equal amount of this antidote to counteract a given amount of poison. So, you presently have the following choice. You can prevent yourself from becoming the murderer of one by giving all 2 cm^3 of the antidote to Valentino or you can prevent Yasmin from becoming the murderer of two by giving 1 cm^3 of the antidote to each of Vladimir and Vasily.

---

[32] We must distinguish between (a) the most fundamental wrong-making feature of an impermissible action and (b) the fact that ultimately grounds the fact that some given act is wrong. The fact that ultimately grounds the fact that it's wrong for you to commit murder even to prevent two others from each committing a comparable murder is that your would-be victim has the kind of value that commands respect. But this fact is not a feature of any action. Rather, it's a feature of your would-be victim. So, in accounting for the wrongness of this *action*, we must appeal to the most fundamental wrong-making feature of wrong actions: that it is the sub-optimal alternative to an act whose prospect you ought to prefer. Of course, you may initially think: "No, what ultimately makes this act wrong is simply that it's an instance of treating someone as a mere means." But it's not that simple. Of course, one of your obligatory ends is that you not treat anyone as a mere means, but the permissibility of your treating someone as a mere means depends on your other obligatory ends and the probabilities that you would achieve them if you were to perform this or some alternative act. Thus, if murdering the one were the only way for you prevent millions of other murders, then it would be permissible for you to do so.

[33] I borrow this sort of case from Emet (2010, 4).

If Kantsequentialism insisted on your having a preference for clean (or cleaner) hands, then it would have to hold that you are required to give all 2 cm^3 of the antidote to Valentino. But Kantsequentialism can instead hold that an agent ought to prefer saving two people to preventing themself from being the murderer of someone they have already treated as a mere means. It's important to note, then, that although there is on Kantsequentialism a constraint against treating someone as a mere means, you are not in *The Antidote* facing the possibility of infringing upon that constraint. For you already used Valentino as a mere means when you injected him with the poison. And there's nothing that you can do now to change that. All you can do at this point is determine both how many people will die and how many murders you will have committed. But there's absolutely no reason why Kantsequentialism can't give priority to the end of minimizing deaths over the end of minimizing the murders that you've committed. Indeed, Kantsequentialism doesn't even have to give you the end of minimizing the murders you've committed. So, despite what Howard claims, there needn't be anything self-centered or self-indulgent about Kantsequentialism's account of constraints.

4.2 Some claim that the teleological approach has counterintuitive implications in intra-agent minimizing cases: Whereas an *inter*-agent minimizing case is one where the agent infringes upon a constraint to prevent more numerous others from infringing upon that same constraint, an *intra*-agent minimizing case is one where the agent infringes upon a constraint to minimize their own infringements of that constraint. Several philosophers (e.g., Brook 1991, Kamm 1996, and Otsuka 2011) have argued that consequentialist views such as ARC have counterintuitive implications in intra-agent minimizing cases. In all their examples, the agent has already treated several people as mere means but can now prevent this treatment from resulting in their being harmed by treating yet another person as a mere means to preventing this harm. These philosophers then assume that the relevant constraint prohibits harming people rather than treating them as mere means. And, so, they conclude that since it is impermissible to harm this other person as a mere means to preventing one's previous actions from causing harm to those

one has already treated as mere means, the rationale for the constraint can't be agent-focused in the way that ARC seems to suggest.[34]

To take just one representative example, consider Kamm's *Guilty Agent Case:* "an agent has set a bomb that will kill five people unless he himself now shoots one other person and places that person's body over the bomb" (1996, 242). Kamm notes that it's counterintuitive to suppose that the agent is permitted to shoot this other person and place his body over the bomb to save the five. But Kamm claims that insofar the consequentialist prohibits us from killing one even to prevent five others from killing only because we would then stand in a particular relationship to the one, a relationship in which we would not stand to the five, the consequentialist should permit us to shoot this other single person. For she claims that "the agent will stand in that same problematic relationship to the five people, if he does not kill the single person" (1996, 242). But this assumes that the problematic relationship is that of 'being one who has killed them' rather than that of 'being one who has treated them as a mere means'. For if it's the latter, then, contrary to Kamm's assertion, the agent doesn't stand in that same problematic relationship to the five people as he does to the one. With respect to the five, he doesn't face the choice of whether to treat them as mere means, for he has already done that and nothing he can do now can change that.[35] By contrast, with respect to the one, he does face the choice of whether to treat them as a mere means to preventing himself from being the murderer of the five. But, on Kantsequentialism, the end of minimizing the instances in which one treats a person as a mere means has priority over other ends, such as that of minimizing the murders

[34] See Brook's case of tossing children into the lion's den (1991, 197), Kamm's "Guilty Agent Case" (1996, 242), and Otsuka's case of the dislodged boulder (2011, 43).

[35] I'm assuming, here, that even if the bomb doesn't go off such that the subjective experiences of the five are left completely unaffected, the five were still *treated* as mere means even if they weren't *used* as mere means. Recall that you treat someone as a means if you behave toward them in a certain way for the sake of realizing some end, intending the presence or participation of some aspect of them to contribute to that end's realization. Thus, the presence or participation of some aspect of them needn't actually contribute to that end's realization. What's more, whether they were treated as mere means seems morally relevant over and above whether this results in their being used as mere means. After all, our most fundamental duty toward persons is to respect them, and treating them as mere means is sufficient to violate this duty.

one has committed.[36] Thus, contrary to what philosophers such as Brook, Kamm, and Otsuka suggest, Kantsequentialism implies that agents should not harm the additional person as a means to preventing harm from coming to those who they have already treated as mere means.

So, if we want to test whether the rationale for constraints is victim-focused or agent-focused, we should look to intra-agent minimizing cases in which the agent hasn't already treated someone as a mere means. Consider, then, the following such case.

> *Intra-Agent Promise-Breaking:* "Three people in Joe's community, Mark, Bill, and Frank, are planning to move over a one week period. Joe has promised each of them that he would help them move. When the time comes to do so, Joe realizes that he cannot keep his promises to all three. He can either keep his promise to Mark and break his promises to Bill and Frank, or he can keep his promises to Bill and Frank and break his promise to Mark." (Lopez et al. 2009, 310)[37]

Intuitively, it seems that Joe should break his promise to Mark so that he can keep his promises to Bill and Frank. Yet, when we consider the *inter*-agent analogue of this case, where Joe must decide whether to break his promise to Mark so that some other agent, David, can keep his promises to Bill and Frank, we get the opposite intuition: Joe should not break his

---

[36] The end of minimizing the instances in which one treats a person as a mere means is a time-neutral one. It's just that, after the guilty agent has already treated the five as a mere means by planting the bomb, the only way for him to minimize the instances in which he treats a person as a mere means is to refrain from treating the sixth person as a mere means.

[37] Another sort of case that we might look at is one where the only way to prevent oneself from treating more numerous people as mere means in the future is to treat one person as a mere means now. The problem is that it's very difficult, if not impossible, to come up with such a case where all the instances of treating as mere means are equally free such that they are morally comparable. For discussions of this issue, see Sturgeon (1996) and Portmore (1998). There are also cases where if you were to refrain from treating X as a mere means today, you *would* then as a matter of fact treat both Y and Z as mere means tomorrow, but you *could* refrain from treating X as a mere means today and then refrain from treating either Y or Z as mere means tomorrow. In this sort of case, I think that you should refrain from treating X as a mere means today because this is entailed by your best option: treating none of X, Y, or Z as mere means. For more on this, see Portmore (2019).

promise to Mark so that David can keep his promises to Bill and Frank.[38] This makes sense only if we adopt Kantsequentialism and take agents to be obligated both to adopt the end of minimizing their own promise-breakings and to give this end priority over that of preventing others from breaking their promises. Thus, on Kantsequentialism, the rationale for why Joe is *permitted* to break his promise to Mark to prevent *himself* from breaking promises to Bill and Frank but is *prohibited* from breaking his promise to Mark to prevent *David* from breaking promises to Bill and Frank is, as Williams (1973, 93–100) has argued, that agents have a special responsibility for their own actions (and promises), a responsibility that they don't have for the actions (or promises) of others. And this is clearly an agent-focused rationale. If, instead, we thought that the rationale was the victim-focused one that lies with the inviolability of persons, then we must instead hold (counterintuitively) that Joe is prohibited from breaking his promise to Mark even to prevent himself from having to break more promises in the future just as he is prohibited from doing so to prevent David from having to break more promises in the future.[39]

Thus, it seems that despite what Brook, Kamm, and Otsuka claim, it's the side-constraint approach, not the teleological approach, that has counterintuitive implications in intra-agent minimizing cases. They were led to the wrong conclusion because they were looking only at cases where the agent had already treated several people as mere means and were wrongly assuming that the relevant constraint must be the one that prohibits harming as opposed to treating as a mere means. But when we look at intra-agent minimizing cases in which no one has yet been treated as a mere means—cases such as *Intra-Agent Promise-Breaking*, we find that it's the side-constraint approach that has counterintuitive implications.

5. Conclusion

I've argued that when we consequentialize constraints by adopting ARC and combining it with

[38] See Lopez et al. (2009).

[39] And we would have to accept this same counterintuitive verdict if we were to hold, as Dougherty (2013) and Setiya (2018) do, that agents are required both to have the end of minimizing promise-breakings-committed-in-order-to-prevent-other-promise-breakings and to give this end priority over that of minimizing promise-breakings overall.

a Kantian theory of value, we arrive at a theory—viz., Kantsequentialism—that has several advantages over the standard side-constraint approach, and it's one that doesn't have any clear disadvantages. If this is right, then, surprisingly, constraints are more plausibly accommodated within a consequentialist framework than within a side-constraint framework, and this is so despite most side-constraint theorists thinking that constraints pose a serious problem for consequentialism.[40]

References

Anderson, E. (1993). *Value in Ethics and Economics*. Cambridge, Mass.: Harvard University Press.

Bader, R. (Forthcoming). "The Dignity of Humanity." In S. Buss and N. Theunissen (eds.), *Rethinking the Value of Humanity*. Oxford: Oxford University Press.

Brook, R. (1991). "Agency and Morality." *Journal of Philosophy* 88: 190–212.

Broome, J. (1991). *Weighing Goods*. New York: Blackwell.

Cox, R. and M. Hammerton (Forthcoming). "Setiya on Consequentialism and Constraints." *Utilitas*.

Cummiskey, D. (1996). *Kantian Consequentialism*. New York: Oxford University Press.

de Marneffe, P. (2013). "Contractualism, Personal Values, and Well-Being." *Social Philosophy and Policy* 30: 51–68.

Dougherty, T. (2016). "Moral Indeterminacy, Normative Powers and Convention." *Ratio* 29: 448–65.

Dougherty, T. (2013). "Agent-Neutral Deontology." *Philosophical Studies* 163: 527–37.

[40] For helpful comments and discussions on earlier drafts, I thank Larry Alexander, Richard Yetter Chappell, Peter de Marneffe, Tom Dougherty, Peter A. Graham, Matthew Hammerton, Liz Harman, Paul Hurley, Chad Lee-Stronach, Jörg Löschke, Daniel Muñoz, and Mark Timmons.

Dreier, J. (1993). "Structures of Normative Theories." *The Monist* 76: 22–40.

Dunaway, B. (2017). "Ethical Vagueness and Practical Reasoning." *The Philosophical Quarterly* 67: 38–60.

Emet, S. F. (2010). "Agent-Relative Restrictions and Agent-Relative Value." *Journal of Ethics & Social Philosophy* 4: 1–13.

Foot, P. (1985). "Utilitarianism and the Virtues." *Mind* 94: 196–209.

Hare, C. (2011). "Obligation and Regret When There is No Fact of the Matter About What Would Have Happened if You Had not Done What You Did." *Noûs* 45: 190–206.

Howard, C. (2021). "Consequentialists Must Kill." *Ethics* 131: 727–53.

Hurley, P. (2018). "Consequentialism and the Standard Story of Action." *The Journal of Ethics* 22: 25–44.

Howard-Snyder, F. (1993). "Rule Consequentialism Is a Rubber Duck." *American Philosophical Quarterly* 30: 271–78.

Jackson, F. and M. Smith (2006). "Absolutist Moral Theories and Uncertainty." *Journal of Philosophy* 103: 267–83.

Kamm, F. M. (1996). *Morality, Mortality. Vol. 2, Rights, Duties, and Status.* New York: Oxford University Press.

Kant, I. (1998). *Groundwork of the Metaphysics of Morals*. Trans. M. Gregor. Cambridge: Cambridge University Press. [Originally published in 1785.]

Kerstein, S. J. (2013). *How to Treat Persons*. Oxford: Oxford University Press.

Korsgaard, C. (1996). *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.

Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.

Löschke, J. (2020). "Value-Based Non-Consequentialism." Habilitationsschrift, Ludwig-Maximilians-Universität München.

Lopez, T., J. Zamzow, M. Gill, and S. Nichols. (2009). "Side Constraints and the Structure of Commonsense Ethics." *Philosophical Perspectives* 23: 305–19.

Muñoz, D. (2021). "The Rejection of Consequentializing." *Journal of Philosophy* 118: 79–96.

Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.

Otsuka, M. (2011). "Are Deontological Constraints Irrational?" In R. M. Bader and J. Meadowcroft (eds.), *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*, pp. 38–58. New York: Cambridge University Press.

Portmore, D. W. (Forthcoming). "Consequentialism." In C. Miller (ed.), *The Bloomsbury Handbook of Ethics, 2nd Edition*. London: Bloomsbury Publishing.

Portmore, D. W. (2019). *Opting for the Best: Oughts and Options*. New York: Oxford University Press.

Portmore, D. W. (2011). *Commonsense Consequentialism: Wherein Morality Meets Rationality*. New York: Oxford University Press.

Portmore, D. W. (1998). "Can Consequentialism Be Reconciled with Our Commonsense Moral Intuitions?" *Philosophical Studies* 91: 1–19.

Ridge, M. (2009). "Consequentialist Kantianism." *Philosophical Perspectives* 23: 421–38.

Scheffler, S. (1985). "Agent-Centred Restrictions, Rationality, and the Virtues." *Mind* 94: 409–19.

Schoenfield, M. (2015). "Moral Vagueness Is Ontic Vagueness." *Ethics* 126: 257–82.

Sen, A. (1983). "Evaluator Relativity and Consequentialist Evaluation." *Philosophy & Public Affairs* 12: 113–32.

Setiya, K. (2018). "Must Consequentialists Kill?" *The Journal of Philosophy* 115: 92–105.

Sinnott-Armstrong, W. (2019). "Consequentialism." In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition. URL = <https://plato.stanford.edu/archives/sum2019/entries/consequentialism/>.

Smith, M. (2009). "Two Kinds of Consequentialism." *Philosophical Issues* 19: 257–72.

Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.

Sturgeon, N. L. (1996). "Anderson on Reason and Value." *Ethics* 106: 509–24.

Thomson, J. J. (2003). *Goodness and Advice*. Princeton, NJ: Princeton University Press.

Thomson, J. J. (1986). "Some Ruminations on Rights." In W. Parent (ed.), *Rights, Restitution, and Risk: Essays in Moral Theory*, pp. 49–65. Cambridge, MA: Harvard University Press.

Vallentyne, P. (1988). "Gimmicky Representations of Moral Theories." *Metaphilosophy* 19: 253–63.

Vessel, J.-P. (2003). "Counterfactuals for Consequentialists." *Philosophical Studies* 112: 103–25.

Williams, B. (1973). "A Critique of Utilitarianism." In J. J. C. Smart and B. Williams (eds.), *Utilitarianism: For and Against*, pp. 75–150. Cambridge: Cambridge University Press.

Williams, J. R. G. (2017). "Indeterminate Oughts." *Ethics* 127: 645–73

Williamson, T. (1994). *Vagueness.* London: Routledge.

Wood, A. W. (2017). "How a Kantian Decides What to Do." In M. C. Altman (ed.), *The Palgrave Kant Handbook*, pp. 263–84. London: Palgrave Macmillan.

Wood, A. W. (1999). *Kant's Ethical Thought*. Cambridge: Cambridge University Press.