

Précis: Commonsense Consequentialism

DOUGLAS W. PORTMORE

Arizona State University

I begin *Commonsense Consequentialism* with a confession: I used to accept utilitarianism. Although there is still something that I find attractive about the view, I have long since given it up, because I couldn't reconcile it with another attractive view: viz., moral rationalism—the view that agents always have decisive reason, all things considered, to fulfill their moral obligations. Utilitarianism conflicts with moral rationalism in part because it requires unreasonable sacrifices. It requires, for instance, that I sacrifice the life of my child if I could thereby bring about even the smallest increase in the net amount of impersonal good. And although the fact that my ϕ -ing would increase the net amount of impersonal good clearly counts as some reason for me to ϕ , I can't help but think that I lack sufficient reason to sacrifice my child's life for the sake of such a small gain in the impersonal good. It seems that the strong reasons that I have to favor the interests of my loved ones and myself decisively oppose the relatively weak reason that I have to slightly increase the amount of impersonal good.

But, as I said, there is something attractive about utilitarianism, something that explains why I was drawn to it in the first place. On reflection, I think it is the idea that how much reason I have to perform an act is a function of how much reason I have to want the world to go in the way that it would go if I were to perform that act. Thus, I hold the somewhat unorthodox view that what is compelling about utilitarianism, and about consequentialism more generally, is the idea that how we ought to act is determined by how we ought to want the world to go. After all, it is through our actions that we affect the way the world goes. Indeed, whenever we face a choice of what to do, we also face a choice of which of various possible worlds to actualize. So what I find compelling is the idea that an agent ought to act so as to actualize the world that, of all the worlds that she could actualize, is the one that she has most reason to want to be actual.

Interestingly, although moral rationalism leads us to reject traditional versions of consequentialism such as utilitarianism, it compels us to accept consequentialism, more generally—which, as I define it, is the view that an act's deontic status is determined by the agent's reasons for and against preferring its outcome to those of the available alternatives. Or at least it does so when conjoined with the teleological conception of practical reasons. That is, it does so via the following argument:

- (1) An act's deontic status is determined by the agent's reasons for and against performing it, such that, if a subject S is morally required to perform an act ϕ , then S has most (indeed, decisive) reason to ϕ . (Moral Rationalism)

- (2) An agent's reasons for and against performing a given act are determined by an agent's reasons for and against preferring its outcome to those of the available alternatives, such that, if S has most reason to ϕ , then, of all the worlds that S could actualize through her various act alternatives, S has most reason to desire w_ϕ , where, for any act x , w_x is the possible world that would be actual if S were to perform x . (The Teleological Conception of Practical Reasons)
- (3) An act's deontic status is determined by an agent's reasons for and against preferring its outcome to those of the available alternatives, such that, if S is morally required to ϕ , then, of all the worlds that S could actualize through her various act alternatives, S has most reason to desire w_ϕ . (Act-Consequentialism)

The book unfolds in two main stages. In the first part of the book, I argue for (3), by arguing for both (1) and (2). That is, I argue for act-consequentialism by arguing for both moral rationalism and the teleological conception of practical reasons, which together entail act-consequentialism. In the second part of the book, I argue that there is a version of act-consequentialism that can accommodate all the central features of commonsense morality: imperfect duties, special obligations, supererogatory acts, agent-centered options, agent-centered restrictions, and the self-other asymmetry. I say "can accommodate," because whether it in fact accommodates these features depends on what the correct substantive account of our reasons for desiring are. Just as utilitarianism can't tell us what to do until we know what utility is, act-consequentialism (as defined above) can't tell us what to do until we know what we have reason to desire. Unfortunately, I have no general substantive account of what we have reason to desire to offer. This is because I don't think that there is any such general theory. For I suspect that there are a plurality of basic considerations that count in favor of desiring various possible worlds. I settle, then, for suggesting that the prospect for act-consequentialism's accommodating our commonsense moral convictions seems promising in that our intuitions about what we have reason to desire appear to be of the sort that when plugged in act-consequentialism yield the commonsense moral verdicts that we might hope to accommodate.¹

The version of act-consequentialism that has the potential to accommodate all the central features of commonsense morality (at least, provided that certain substantive views about what we have reason to desire are correct) is what I call *commonsense consequentialism*. I'll state the view in a moment, but first I need to explain some terminology. Options are just acts that an agent can perform. Some options are more specific than others in the sense that performing some logically necessitates performing others. For instance, raising my hand quickly logically necessitates raising my hand and so is more specific than it. And raising my hand and then waving it about above my

¹ For some examples, see Douglas W. Portmore, *Commonsense Consequentialism: Wherein morality meets rationality*. New York: Oxford University Press, 2011, pp. 114–115.

head logically necessitates both raising my hand and waving it about above my head and so is more specific than either. Some of my options will even be maximally specific; these are the options whose performance is not logically necessitated by the performance of any other option. I call these *maximal options*. And I call all other options *non-maximal options*. Given this terminology and simplifying a bit, commonsense consequentialism can be stated as follows:

- CSC (a) *S*'s performing a maximal option M_1 is permissible if and only, and because, there is no alternative maximal option M_2 such that *S* has both more moral reason and more reason, all things considered, to desire w_{M_2} than to desire w_{M_1} .
- (b) *S*'s performing a non-maximal option N_i is permissible if and only if, and because, there is some permissible maximal option M_i such that performing M_i logically necessitates performing N_i .

CSC has a number of unique features. To help understand them better, it will be useful to compare it to *traditional act-consequentialism*:

- TAC *S*'s performing an option ϕ is morally permissible if and only if, and because, there is no alternative option ψ such that w_ψ is better (impersonally speaking) than w_ϕ .

I'll limit my discussion here to three key differences between the two views, the first being that whereas TAC evaluates both maximal and non-maximal options on the basis of their outcomes, CSC does not. CSC evaluates maximal options on the basis of their outcomes. But it evaluates non-maximal options on the basis of whether their performance is logically necessitated by the performance of some permissible maximal option. This allows CSC to avoid a problem that arises for TAC. The problem arises in the following sort of case. Suppose that Professor Horrible gets an invitation to review a book. And let's assume that if he were to intend both to accept the invitation and to write the review, then that's precisely what he would do, which is what would be best. Let's assume, then, that his accepting and writing is an option. Moreover, let's assume that all his best maximal options involve his accepting and writing. Unfortunately, though, Professor Horrible is, as his name suggests, a horrible person who intends to accept and not write, which is the worse thing that he could do. Given this intention, it is clear that Professor Horrible would not write the review if he were to accept the invitation to do so. So accepting the invitation would have bad consequences. And, thus, TAC implies that Professor Horrible should not accept the invitation. Yet TAC implies that Professor Horrible should accept and write, for performing this conjunctive act would have excellent consequences. So TAC implies both that he should accept and write and that he should not accept. For reasons that I explain in the book, we should be

loath to accept such a view.² We should instead hold that if an agent should perform a conjunctive act, then she should perform each of its conjuncts. CSC ensures that this will be the case by holding that whether an agent should perform any individual act (that is, any non-maximal option) just depends on whether she should perform some conjunctive act (that is, some maximal option) of which it is a part.

The second key difference between the two views is that whereas TAC ranks outcomes along a single evaluative dimension (that is, their impersonal goodness), CSC ranks outcomes according to a function of two auxiliary rankings, each along a different evaluative dimension—one being a ranking in terms how much *moral* reason the agent has to desire them and the other being a ranking in terms of how much reason, *all things considered*, the agent has to desire them. The reason for this dual-ranking structure lies with the fact that non-moral reasons can, and sometimes do, prevent moral reasons from generating moral requirements, or so I argue. Thus, non-moral reasons serve to limit what we can be morally required to do. If, for instance, the non-moral reasons that you have to perform some self-interested act, *x*, outweigh the moral reasons that you have to perform some altruistic alternative act, *y*, then you cannot be morally required to sacrifice your self-interest and perform *y* as opposed to *x*. Of course, you would not be morally required to perform *x* either, for you would also be morally permitted to do what you have more moral reason to do rather than what you have less moral reason to do.

The upshot of all this is that the deontic status of an action is a function of both moral reasons and non-moral reasons. That function is, as I argue in chapter 5, given by the following *meta*-criterion of rightness:

META S's performing ϕ is morally permissible if and only if, and because, there is no available alternative, ψ , that S has both more moral reason and more reason, all things considered, to perform.

And, in the same chapter, I argue that the only way to accommodate META within a consequentialist framework is to hold that one outcome outranks another, on the agent's evaluative ranking, if and only if she has both more moral reason and more reason, all things considered, to want the one to obtain than to want the other to obtain. Thus, on CSC, the ranking of outcomes is a function of two auxiliary rankings: one a ranking in terms of how much moral reason the agent has to desire that they obtain and the other in terms of how much reason, all things considered, the agent has to desire that they obtain.

This dual-ranking structure is what allows CSC to accommodate agent-centered options (of both the agent-favoring and the agent-sacrificing varieties), for it provides agents with the option to give their interests more (as in the case of agent-favoring options) or less (as in the case of agent-sacrificing options) weight than the equivalent interests of others. TAC, by contrast, ranks outcomes along a single evaluative

² See *Commonsense Consequentialism*, especially pp. 180–182.

dimension: their impersonal value. Unlike CSC, then, there is no way for TAC to accommodate options to give one's own interests anything but the weight that they have from the impersonal perspective.

To account for the agent-favoring option that S has, say, to perform presently either an altruistic act or a self-interested act, the proponent of CSC need only claim that there is both a morally permissible maximal option in which she presently performs an altruistic act and a morally permissible maximal option in which she presently performs a self-interested act. Perhaps, as I argue, what matters is not whether S performs an altruistic or self-interested act presently, but only whether she performs some maximal option that strikes a reasonable balance over time between both types of acts.

To account for the agent-sacrificing option that S has, say, to use her last dose of painkillers so as to alleviate either some stranger's mild headache or her own more severe headache, the proponent of CSC need only claim that, other things being equal, she has more moral reason to prefer the outcome of a maximal option in which she alleviates the stranger's headache, but more reason, all things considered, to prefer the outcome of an otherwise identical maximal option in which she instead alleviates her own more severe headache.

The third and perhaps most crucial difference between the two views is that whereas TAC ranks outcomes according to their impersonal goodness, CSC ranks outcomes according to how much reason the agent has to desire that they obtain. To illustrate the difference, suppose that I can effect only one of two possible worlds: either the one in which I save my own daughter (call this w_1) or the one in which I ensure that some stranger saves his slightly more gifted daughter (call this w_2). And let me just stipulate that w_2 is slightly better than w_1 , as the stranger's daughter would contribute slightly more impersonal good to the world than my daughter would. We will assume, though, that everything else is equal. Now, despite the fact that w_2 is slightly better than w_1 , it seems that I have more reason to desire that w_1 be actualized than to desire that w_2 be actualized, for it seems that the reason that I have to desire the world in which my daughter is saved outweighs the reason that I have to desire the world in which there is slightly more impersonal good. And it is not just that I have more reason, *all things considered*, to desire that w_1 be actualized. Given the moral importance of the special relationship that I have with my daughter, I also have more *moral* reason to desire that w_1 be actualized. Indeed, if I did not care more about the preservation of my daughter's life than I did about such small increases in the impersonal good, I would be open to moral criticism. Parents ought, morally speaking, to care more about the preservation of their own children than they do about such small increases in the impersonal good.

If all this is right, then CSC and TAC will sometimes rank these two worlds differently. On TAC, there is only one possible ranking of these two worlds, and, on that ranking, w_2 outranks w_1 . On CSC, however, there is potentially a different ranking for each agent, as each agent is to rank outcomes according to how much reason she has to desire that they obtain. Since one agent can have more reason than another to desire that a certain outcome obtains, different agents can end up with different rankings. This is true in the above example. On my evaluative ranking, w_1 outranks w_2 , but, on the

stranger's (or on an impartial bystander's) evaluative ranking, w_2 outranks w_1 . Thus, whereas TAC holds that we all have the same goal of bringing about the best possible outcome, CSC holds that each individual agent has the potentially distinct goal of bringing about the outcome that *she* has most reason to desire.

The problem with TAC's ranking outcomes according to their value, I believe, is that it excludes from consideration all agent-relative reasons. This is because the impersonal value of a possible world correlates with only a subset of an agent's reasons for desiring its actualization: her agent-neutral reasons. But agents can have agent-relative reasons for desiring its actualization, as is the case with my reasons for desiring w_1 —the possible world in which my daughter is saved. Moreover, these agent-relative reasons can tip the balance such that agents sometimes have better overall reason to prefer a worse world to a better world, as is the case where I have better overall reason to prefer w_1 to w_2 . And, as the teleological conception of practical reasons implies, if I have better overall reason to *prefer* w_1 to w_2 , then I have better overall reason to *do* what will actualize w_1 as opposed to what will actualize w_2 . Given this and moral rationalism, we should conclude that, contrary to what TAC implies, I cannot be obligated to act so as to actualize w_2 as opposed to w_1 even though w_2 has more impersonal value. The lesson, then, is that if we want to take into account all of our reasons for desiring a given outcome and, thereby, all of our reasons for acting so as to produce that outcome, we must rank outcomes according to how much reason the agent has to desire that they obtain as opposed to how valuable they are.

It is because CSC ranks outcomes/worlds according to how much reason the agent has to desire their actualization that it, unlike TAC, can accommodate features such as special obligations and agent-centered restrictions. To account, for instance, for the special obligation that I have to save my own daughter in the case from above where the only other option is to ensure that some stranger saves his slightly more gifted daughter, the proponent of CSC need only claim that, other things being equal, I have both more moral reason and more reason, all things considered, to desire that w_1 be actualized than to desire that w_2 be actualized. To account for an agent-centered restriction against, say, the commission of murder, the proponent of CSC need only claim that, other things being equal, S has both more moral reason and more reason, all things considered, to prefer the outcome of a maximal option in which she refrains from murdering some individual to the outcome of an otherwise identical maximal option in which she instead murders that individual so as to prevent less than n others from each committing a comparable murder—where n marks the threshold for this non-absolute restriction.³

From this brief summary, I hope that the reader will have gained a sense of why I think that we should accept act-consequentialism and why I think that a version of act-consequentialism (*viz.*, CSC) has the potential to accommodate the central features of

³ Note the ease with which commonsense consequentialism can account for the fact that agent-centered restrictions have thresholds: at some point, the agent-neutral reason that one has to want there to be fewer infringements of the restriction overall outweighs the agent-relative reason that one has to want there to be no personal infringements of that restriction. Deontological views have a much more difficult time accounting for thresholds.

commonsense morality. That said, the aim of the book is not convince anyone that commonsense consequentialism is the most plausible moral theory. If that were its aim, I would have needed to provide a more fully developed account of what agents have reason to desire. What's more, I would have needed to survey all of CSC's rivals and made the relevant comparative assessments. I have done neither. But, then, the aim of the book is considerably more modest. The aim is only to have developed the most plausible version of a particular breed of act-consequentialism, one that makes the deontic statuses of actions a function of the agent's reasons for desiring their outcomes. The advantages and disadvantages of theories that take the deontic statuses of actions to be a function of the value of their outcomes have been well explored, but the same cannot be said of the advantages and disadvantages of theories that take the deontic statuses of actions to be a function of how much reason agents have to want their outcomes to obtain. My hope, though, is to have gotten this exploration well underway and to have uncovered enough to demonstrate that such theories merit further consideration.

Replies to Gert, Hurley, and Tenenbaum

DOUGLAS W. PORTMORE

Arizona State University

I am very grateful to Josh Gert, Paul Hurley, and Sergio Tenenbaum for their challenging criticisms. I am also grateful that they have each chosen to focus on core parts of the book. Unfortunately, I won't have space below to address all of their criticisms. Even with respect to those that I do address, I won't have space to give them their due. Nevertheless, I have tried to address their core objections as best as I can.

Reply to Gert

Moral rationalism is the view that, as a matter of conceptual necessity, no agent can have sufficient reason to act wrongly. Since *moral rationalism* (MR) is a conceptual thesis, I argue for it on the basis of two other conceptual theses. One holds that *wrongdoing entails blameworthiness* (WEB) and the other holds that *blameworthiness entails the lack of sufficient reason* (BELS). The argument is this:

WEB S's ϕ -ing is wrong only if S would be blameworthy for freely and knowledgeably ϕ -ing.

BELS S would be blameworthy for freely and knowledgeably ϕ -ing only if S

does not have sufficient reason to ϕ .

Therefore,

MR S's ϕ -ing is wrong only if S does not have sufficient reason to ϕ .

The point of the qualifier "freely and knowledgeably" is to allow that an agent can do wrong and yet be exculpated from blame either because her act was not free or because she lacked certain relevant knowledge. You should assume, then, that all the actions that I discuss below are done freely and knowledgeably.

I do not argue for WEB. I merely cite the fact that it is intuitively plausible and, consequently, widely endorsed. I do, however, argue for BELS. First, I argue that we can hold an agent morally responsible (and, thus, potentially blameworthy) for her actions only if she has the capacity to respond appropriately to the relevant reasons. Second, I claim that, in flawlessly exercising this capacity, an agent could be led to perform any action that she has sufficient reason to perform. So if, contrary to BELS, an agent could be blameworthy for performing an action that she had sufficient reason to perform, then the very capacity that opens the door to her being blameworthy is the one that leads her to perform the blameworthy act. But, "surely, it cannot be that the very capacity that opens the door to an agent's being blameworthy is the one that leads her to perform blameworthy acts" (48).⁴ After all, there is something quite strange about holding an agent morally responsible for having performed an immoral action (and, hence, blameworthy) in virtue of her having a certain capacity (as well as her meeting various other necessary conditions for moral responsibility) when she was led, in flawlessly exercising it, to perform that action—the action for which she is now blameworthy. Thus, we should accept BELS so as to avoid being committed to such a strange view.

This argument for BELS is based on the intuition expressed by the quote just above. Gert calls it "Portmore's strong intuition." I'll abbreviate this as PSI, and I'll give a more precise formulation of it below. Gert rejects my argument for MR, because he rejects BELS. And he rejects my argument for BELS, because he rejects PSI. He also denies that BELS is, as I claim, plausible on its face. He says,

I...do not think BELS is in fact very plausible on its face.... Against its intuitive plausibility, it seems to me that many standard cases of immoral action provide counterexamples. For example, cheating on the last exam of one's college career because otherwise one will fail the class and have to enroll another semester is morally wrong, but I do not think one is *rationally* required to refrain.

But note that such putative counterexamples are, if anything, counterexamples to MR, not to BELS. For they are examples in which, supposedly, it would be wrong for the agent to do something that she has sufficient reason to do—e.g., to cheat on an exam. It is, nevertheless, important for me to address such putative counterexamples, for,

⁴ All parenthetical page references within the body of the paper are to *Commonsense Consequentialism*.

ultimately, it is MR that I'm defending and that Gert is disputing. Unfortunately, though, I think that Gert and I just have conflicting intuitions. I can imagine situations in which one has sufficient reason to cheat, such as those in which not cheating would have disastrous consequences for oneself or others. But in all those situations it seems morally permissible to cheat. We certainly think that it would be permissible to cheat if someone's life were at stake. But Gert thinks that there can be sufficient reason to cheat even if very little is at stake, such as where one just doesn't want to endure another semester of college. And neither of us thinks that it's permissible to cheat if that's all that is at stake.

The reason that Gert thinks that there is sufficient reason to cheat even when there is very little at stake is that he believes that moral reasons have little to no rational requiring strength. He holds, for instance, that the moral reason that one has to keep one's promise to hold office hours has so little rational requiring strength that there would be sufficient reason to abandon one's office hours "for no reason at all beyond the fact that one didn't feel like holding them." What's more, Gert thinks that altruistic moral reasons can "*never*, in themselves, rationally require action."⁵ He even thinks that the moral reason that one has to save forty children from serious malnutrition has "no (or not much) requiring strength."⁶ But if like me, and unlike Gert, you think that such moral reasons have significant rational requiring strength such that you can be *rationally* (and not just morally) required to keep a promise, save forty children, or refrain from cheating even when you would be slightly better off doing otherwise, then you needn't think that "standard cases of immoral action," such as cheating on an exam because you prefer not to endure another semester of college, provide counterexamples to MR.

Besides citing putative counterexamples against BELS (or MR), Gert tries to undermine my argument for BELS by providing a putative counterexample to one of the crucial premises in my argument for BELS: viz., PSI, which holds that, for any capacity C, if C opens the door to an agent's being blameworthy, then it is not the case that, in flawlessly exercising C, one could be led to perform a blameworthy act. Gert claims that PSI is false, because "the capacity to choose between right and wrong" is one that opens the door to an agent's being blameworthy and yet is also one that could, even when exercised flawlessly, result in one's performing a blameworthy action. Now, you might think that in flawlessly exercising the capacity to choose between right and wrong one would necessarily be led to choose to do right. But Gert insists that, "in flawlessly exercising the capacity to choose between right and wrong, one can choose wrong." This is because, as Gert conceives of it, the capacity to choose between right and wrong is, as he puts it, "essentially a volitional capacity," ...one that "involves the capacity to distinguish right from wrong, but goes beyond it to include the capacity [to] form an effective intention to perform either sort of act: a right one or a wrong one."⁷ To

⁵ Joshua Gert, *Brute Rationality: Normativity and Human Action*. Cambridge: Cambridge University Press, 2004, p. 89.

⁶ Gert, *Brute Rationality*, p. 92.

⁷ Quoted with permission from an email dated 8/27/13.

understand this better, consider the following example. Suppose that, despite knowing that it would be wrong to do so, I want to kill Smith's child in revenge for her having killed my child. But suppose that when the opportunity to kill the child arises, I just can't bring myself to do it. This, according to Gert, would be a flawed exercise of the capacity to choose between right and wrong—an instance of weakness of the will, as he conceives of it. Only if I had been able to bring myself to do as I wanted and kill the child would I have flawlessly exercised the capacity to choose between right and wrong.⁸

Admittedly, it does seem that having the capacity to choose between right and wrong—one that entails having the ability both to differentiate between right and wrong and to form effective intentions—is necessary for (and, thus, in some sense opens the door to) one's being blameworthy. Moreover, it's clear from the above example that one can flawlessly exercise this capacity and end up choosing to do wrong—choosing to kill Smith's child, for instance. But note that, in the case of the capacity to choose between right and wrong, one is not *led* to choose to do wrong (e.g., to kill Smith's child) as a result of exercising it flawlessly. Rather, what leads one to choose to do wrong (to choose to kill Smith's child) is one's ignoble end—e.g., a desire to exact revenge by killing an innocent child. In flawlessly exercising the capacity to choose between right and wrong, one is led (1) to identify killing Smith's child as wrong, (2) to identify refraining from killing Smith's child as right, and (3) to form an effective intention to perform whichever act (right or wrong) that one chooses. But one is not led to choose wrong in flawlessly exercising this capacity.

Importantly, the capacity to respond appropriately to the relevant reasons is different in this respect. In flawlessly exercising this capacity, one is led both (i) to choose to perform some act that one has sufficient reason to perform and (ii) to choose arbitrarily between all the acts that one has sufficient reason to perform—at least, whenever there is no more reason to perform any one of these acts than there is to perform any other. So assume for reductio that one can have sufficient reason to do wrong. If such is possible, it seems that this will often be the case because one has no more reason to do right than to do wrong—that is, no more reason to choose to perform R than to choose to perform W. Seeing that this is the case and exercising flawlessly the capacity to respond appropriately to the relevant reasons, one would then be led to choose arbitrarily between R and W by, say, flipping a coin. And depending on how the coin toss goes, one will sometimes be led to choose to perform W. So we see that, if we can have sufficient reason to do wrong, we will sometimes be led to do wrong in flawlessly exercising our capacity to respond appropriately to the relevant reasons—a capacity that opens the door to our being culpable for any wrongdoing we commit. And this seems deeply problematic.

So Gert's capacity to choose between right and wrong fails as a counterexample to PSI. To succeed, it would have to be a capacity that both opens the door to one's being

⁸ I confirmed with Gert via email (dated 8/27/13) that this is indeed an apt example of failing to flawlessly exercise the capacity to choose between right and wrong.

blameworthy and is such that it could lead one to choose to do wrong. But Gert's capacity is not one that could lead one choose to do wrong. Indeed, it is, perhaps, telling that Gert does not say that, in flawlessly exercising the capacity to choose between right and wrong, one is *led* to choose wrong. Instead, he says only that, "in flawlessly exercising the capacity to choose between right and wrong, one *can* choose wrong [emphasis mine]." So I'm not convinced by Gert's putative counterexample to PSI. And if I'm right about PSI's being true, then I can establish 2.19 (an instance of PSI) via the more general PSI. And if that's right, then I do not, as Gert claims, need to beg any questions against those who deny MR in asserting 2.19. I can, as I did in the book, just appeal to PSI in defense of 2.19, which supports 2.19 without presupposing MR. I can then use 2.19 to defend BELS, and in turn use BELS to defend MR, without begging the question.

Reply to Hurley

Besides moral rationalism (MR), the other claim that I rely on in arguing for my particular breed of act-consequentialism is the *teleological conception* of (practical) *reasons*:

TCR For any two act-alternatives ϕ and ψ , S has more reason to ϕ than to ψ if and only if, and because, S has more reason to desire w_ϕ than to desire w_ψ , where w_ϕ is the world that would be actual if S were to ϕ and w_ψ is the world that would be actual if S were to ψ .

Hurley takes issue with TCR and with my argument for it. Hurley interprets me as offering support for TCR by appealing both to "a state of affairs centered account of desire, upon which all desires are that some state of affairs obtains" and to a state-of-affairs-centered account of action, upon which all actions aim at bringing about some state of affairs. As Hurley notes, these state-of-affairs-centered accounts of desire and action can seem problematic. For sometimes what we desire and aim at is not some state of affairs, but an activity. Sometimes, for instance, I just desire to run. Even so, we might think that my desiring *to run* is equivalent to my desiring *that I run*. But Hurley would deny this, arguing, perhaps, that only a desire *that I run*, and not a desire *to run*, could ever motivate me, say, to go out this morning and hire a personal trainer who would then force me to run this evening.⁹ But even if we think that my desiring *to run* is not equivalent to my desiring *that I will run this evening* in that only the latter could motivate me to hire a personal trainer, we might still think that my desiring *to run presently* is equivalent to my desiring *that I run presently*.

It's clear, though, that Hurley thinks that there is an important difference between desiring/aiming to run presently and desiring/aiming to bring it about that I run presently. I'm not so sure. But, in the end, I don't think that the issue is important to

⁹ Actually, Howard Nye suggested this argument to me.

our appraisal of TCR or my argument for it. Admittedly, I did at one point in the book suggest that there is no important distinction between my desiring/aiming to ϕ at t and my desiring/aiming that I ϕ at t . After all, I did say,

It is through our actions that we affect the way the world goes. Indeed, whenever we face a choice of what to do, we also face a choice of which of various possible worlds to actualize. Moreover, whenever we act *intentionally*, we act with the aim of making the world go a certain way. The aim need not be anything having to do with the causal consequences of the act. The aim could be nothing more than to bring it about that one performs the act. For instance, one could intend to run merely for the sake of bringing it about that one runs. The fact remains, though, that for every intentional action there is some end at which the agent aims. (56)

Nevertheless, it seems to me that I can remain neutral on this issue. Because whether or not my desiring/aiming to ϕ at t is equivalent to my desiring/aiming to bring it about that I ϕ at t , it is absolutely clear and “non-controversial” even by Hurley’s own lights that if I ϕ at t , the actual world will be the one in which I have ϕ -ed at t (viz., w_ϕ) and not the one in which I have ψ -ed at t (viz., w_ψ), assuming, as I am, that ϕ -ing and ψ -ing are mutually incompatible alternatives. And, as a proponent of TCR, I am committed to one thing only, and that’s not a state-of-affairs-centered account of desire or action. Rather, it’s that I have more reason to ϕ than to ψ if and only if, and because, I have more reason to desire w_ϕ than to desire w_ψ . Whether in ϕ -ing, I actually desired or aimed to bring it about that w_ϕ obtains is not important for my purposes.

So Hurley rightly claims that “an appeal to the state of affairs centered nature of action does not seem to provide support for TCR, absent additional argument.” And he claims that the same holds for an appeal to a state-of-affairs-centered account of desire. On both counts, I agree entirely. But the above-quoted paragraph was not offered as an argument for TCR. Rather, it was just as a suggestion as to why one *might* find TCR attractive in the first place. It’s important, then, that Hurley goes on to claim: “For those non-consequentialists who are as skeptical about Portmore’s state-of-affairs-centered accounts of action and desire as they are about his state-of-affairs centered account of reason, his positive arguments for TCR will also prove unpersuasive.” This is what’s crucial, because my intention was never to commit myself to a state-of-affairs-centered account of either action or desire, but only to TCR and the positive arguments that I give in support of it. What’s more, I had always thought that my arguments for TCR didn’t hinge on the correctness of any particular account of action or desire. So let’s see why Hurley thinks otherwise.

Let’s start, though, by looking at my argument, which has two main parts. First, I argue that there is a correlation between an agent’s reasons for actions and her reasons for desiring the worlds that result from performing them such that, for any two act-alternatives ϕ and ψ , S has more reason to ϕ than to ψ if and only if S has more reason to desire w_ϕ than to desire w_ψ . (Note the absence of “and because” after “if and only if.”) Call this correlation thesis ‘CT’. Second, I argue that TCR offers the most plausible explanation for CT. Now, Hurley offers no objection to the first part of my argument—that is, he offers no objection to CT or my argument for it. Rather, he objects to the

second part of my argument—that is, he rejects my arguments for TCR’s offering the most plausible explanation for CT.

There are, in fact, four possible explanations for CT:

- E1 For any two act-alternatives ϕ and ψ , if S has more reason to ϕ than to ψ , this is because S has more reason to desire w_ϕ than to desire w_ψ .
- E2 For any two act-alternatives ϕ and ψ , if S has more reason to desire w_ϕ than to desire w_ψ , this is because S has more reason to ϕ than to ψ .
- E3 For any two act-alternatives ϕ and ψ , if S has both more reason to ϕ than to ψ and more reason to desire w_ϕ than to desire w_ψ , this is because p , where ‘ p ’ specifies some shared explanation such as ‘ w_ϕ is better than w_ψ ’.
- E4 For *certain* act-alternatives ϕ and ψ , S has more reason to ϕ than to ψ , because S has more reason to desire w_ϕ than to desire w_ψ . For all *other* act-alternatives α and β , though, either (1) S has more reason to desire w_α than to desire w_β , because S has more reason to α than to β or (2) S has both more reason to α than to β and more reason to desire w_α than to desire w_β , because p , where ‘ p ’ specifies some shared explanation such as ‘ w_α is better than w_β ’.

TCR is committed to E1’s being the correct explanation, and so I defend TCR by arguing that E1 is more plausible than any of its three alternatives: E2-E4. Against E2, I offer the following counterexample. Suppose that I’m choosing which savings account to put my money in. Let a_1 be the act of putting my money into savings account SA₁, which earns 1% annually, and let a_2 be the act of putting my money into savings account SA₂, which earns 2% annually. Let w_1 be the world in which I put my money into SA₁ and consequently receive a 1% annual return on my investment, and let w_2 be the world in which I put my money into SA₂ and consequently receive a 2% annual return on my investment. Now, it just seems obvious to me that I have more reason to perform a_2 than to perform a_1 and that this is because I have more reasons to desire w_2 than to desire w_1 . And if this is right, then E2 is false. For E2 makes a universal claim for which there is a clear counterexample.

Against E3, I have nothing decisive to say. But it seems to me that the most obvious candidates for ‘ p ’ can’t be right. It can’t be, for instance, that ‘ p ’ stands for ‘ w_ϕ is better than w_ψ ’, because sometimes I have more reason to desire w_ψ than to desire w_ϕ even though w_ϕ is better than w_ψ . To see this, let w_ϕ be the world in which I use my only life preserver to save my own child rather than some stranger’s slightly more gifted child, and let w_ψ be the world in which I use my only life preserver to save some stranger’s slightly more gifted child rather than my own child. Even if it would be better if the more gifted child were saved, I still have more reason to want it to be my child who is saved. Also, it won’t do to let ‘ p ’ stand for ‘ w_ϕ is better for the agent than w_ψ is’, because sometimes I have more reason to desire w_ψ than to desire w_ϕ even though w_ϕ is better for me than w_ψ is. It seems, for instance, that I can have more reason to desire

the world in which I die and my child lives than to desire the world in which I live and my child dies even if the latter would be better for me. So neither of these two most obvious candidates for 'p' is adequate. And I'm at loss as to what a more plausible candidate for 'p' would be.

This leaves us with E1 and E4. E1 is the more unified of the two and so offers greater explanatory power. Of course, if there were any obvious counterexamples to E1, then we should reject E1 despite its greater explanatory power (just as we rejected E2 despite its greater explanatory power). But I don't think that there are any clear counterexamples to E1. Hurley would disagree. He says that it seems backwards to think that "my reasons to fulfill my promise are fundamentally my reasons to bring it about that my promise is fulfilled." But does it seem backwards to claim that because people have a certain kind of value, we have a reason to want our promises to them to be fulfilled and, thus, a reason to act so as to bring it about that our promises to them are fulfilled? It doesn't seem backwards to me. Moreover, as I argue in chapter 4 of my book, explaining my reasons to refrain from breaking my promises in terms of the reasons that I have to prefer, other things being equal, the world in which I break as few of them as possible has the added advantage of allowing us to explain why, if I have made three comparably important promises—specifically, to do x at t_1 , y at t_2 , and z at t_3 , it is permissible for me to refrain from doing x at t_1 if this is the only way for me to ensure that I will do both y at t_2 and z at t_3 . At least, this way I'll break only one of my promises, and the situation, I'm stipulating, is one where the choice is between, on the one hand, breaking just one of my promises and, on the other hand, breaking two or three of them.

So Hurley has different intuitions. He thinks that it is backwards to think my reasons to fulfill my promises derive from my reasons to bring it about that my promises are fulfilled. So, perhaps, he thinks that E4 is more plausible than E1. That is, perhaps, he thinks that when it comes to the savings account case, it is plausible to think that I have more reason to perform a_2 than to perform a_1 , because I have more reasons to desire w_2 than to desire w_1 . But when it comes to promises, the explanatory direction reverses: I have more reason to desire the world in which my promises are kept because I have more reason to keep my promises. This comes at a cost in that E4 has less explanatory power than E1, but perhaps it seems to Hurley to be less costly than adopting E1 given how backwards he thinks it gets things when it comes to cases involving promises.

Hurley could instead go with E2 and say that I have things backwards even in the savings account case. He says,

But with the rejection of the state of affairs centered account of value kept clearly in view, such cases seem to pose no problem for opponents of TCR. They can readily allow both that the value of states of affairs, e.g. as better for me or better overall, provides me with reasons to act, in particular to promote, and that the value of persons and things provides me with other reasons to act – to treat people respectfully, preserve valuable things, etc. ...In short, it is unclear why such examples cannot readily be accommodated by the opponent of TCR in an account upon which it is the left (reasons to do) side of the bi-conditional [i.e., CT] that in the relevant sense has explanatory priority.

Hurley is right that the opponent of TCR can explain CT by appealing to E2, insisting, then, that even in the savings account case it is the fact that I have more reason to perform a_2 than to perform a_1 that explains the fact that I have more reasons to desire w_2 than to desire w_1 . I, of course, don't deny that the opponent TCR can make this move. I just deny that it is a plausible one. For it seems counterintuitive to suppose that reasons for action have explanatory priority in the savings account case.

I also don't deny that the opponent of TCR can appeal both to the value of states of affairs and to the value of persons in explaining why we should promote valuable states of affairs and act respectfully to persons. Interestingly, the proponent of TCR can do so as well. It's just that the proponent of TCR would, first, appeal both to the value of states of affairs and to the value of persons in explaining why we have reasons to desire the worlds in which we promote valuable states of affairs and act respectfully to persons, and then, second, appeal to this fact in explaining why we have reason to promote valuable states of affairs and to act respectfully to persons.

Reply to Tenenbaum

Tenenbaum focuses his criticisms on chapter 4 and my consequentializing project. In the chapters leading up to chapter 4, I argue for both MR and TCR, which together entail act-consequentialism:

AC An act's deontic status is determined by an agent's reasons for and against preferring its outcome to those of the available alternatives, such that, if S is morally required to ϕ , then, of all the worlds that S could actualize through her various act alternatives, S has most reason to desire w_ϕ .¹⁰

Then, in chapter 4, I argue that AC can potentially do at least as good a job of accommodating our considered moral convictions as the best nonconsequentialist theories can. For as I demonstrate there is, for every remotely plausible nonconsequentialist theory, an AC-counterpart theory that yields in every possible world the exact same deontic verdicts that it yields. Wherever the nonconsequentialist theory holds that S ought not to ϕ in virtue of consideration C, the AC-counterpart theory holds instead that S ought to prefer $w_{-\phi}$ (the world in which she doesn't ϕ) to w_ϕ (the world in which she ϕ s) in virtue of C and that, in virtue of this fact, S ought not to ϕ . In coming up with this AC-counterpart theory, we have consequentialized the nonconsequentialist theory with which we started.

¹⁰ Actually, for this entailment to go through, I need something stronger than MR, something that also posits an explanatory relation between the deontic statuses of actions and the reasons to perform them. But I assume that proponents of MR would be happy to posit such a relation. See *Commonsense Consequentialism*, note 19 on pp. 33-34.

“The motive for consequentializing is...to arrive at a moral theory that comports with our considered moral convictions without having to forfeit, as nonconsequentialist theories do, the compelling idea that an act’s deontic status is determined by the agent’s reasons for preferring its outcome to those of the available alternatives such that it can never be morally wrong for her to act so as to bring about the outcome that she has most reason to want to obtain” (8). So the problem with nonconsequentialist theories is that they have to forfeit (act-)consequentialism, which we are driven to accept via MR and TCR. And the problem with many traditional consequentialist theories is that they don’t comport with our considered moral convictions. By consequentializing, though, the hope is that we can get the best of both worlds.

The consequentializing project does not involve simply taking whatever set of deontic verdicts or nonconsequentialist theory that one likes and then consequentializing it. Nor does it involve simply taking whatever views one has about which worlds agents have most reason to desire, plugging that into the following consequentialist bi-conditional, and hoping for the best—that is, hoping that the resultant theory doesn’t have any counterintuitive implications.

BC S’s ϕ -ing is morally permissible if and only if there is no alternative ψ such that S has both more moral reason and more reason, all things considered, to desire w_ψ than to desire w_ϕ .

The project is much more nuanced than either of these. For although the process involves holding BC constant, it involves some give and take both with respect to what we think agents ought to do and with respect to which worlds we think that they have most reason (or most moral reason) to desire. Sometimes, we might, on reflection, change our minds about what agents ought to do given some entrenched view about which worlds agents have most reason to desire and what that, in conjunction with BC, implies about what they ought to do. But other times we might change our initial view about which worlds agents have most reason to desire, because plugging that into BC yields counterintuitive verdicts regarding what they ought to do. At some point, though, we will reach reflective equilibrium. And the hope is that the resulting theory will have all the merits of consequentialism but, unlike traditional consequentialist theories, will have few, if any, counterintuitive implications. Moreover, the hope is that whatever counterintuitive implications it does have will seem, on reflection, worth countenancing given the intuitive plausibility of my arguments for TCR and MR and their entailment of act-consequentialism.

As far as I can tell, there are only three possible ways that this project could fail:

- A. The resulting theory has unacceptable implications. That is, its implications are sufficiently counterintuitive that it would be better to reject either MR or TCR than to accept these implications.
- B. The resulting theory has acceptable implications, but, as it turns out, it is merely a notational variant of some nonconsequentialist theory.

- C. The resulting theory has acceptable implications, but, as it turns out, it is not consequentialist, because it is committed to explaining why an agent ought to prefer $w_{\neg\phi}$ to w_{ϕ} by appealing to the fact that she ought not to ϕ .

My (A)-(C) are meant to roughly correspond to Tenenbaum's trilemma involving his (1)-(3). Tenenbaum views this as a trilemma, because he thinks that the project is doomed to fail in at least one of these three ways. He thinks, for instance, that if it doesn't fail in way (1), then it will fail in either way (2) or (3). And he thinks that if it doesn't fail in either way (2) or (3), it will fail in way (1). I assume that Tenenbaum would say the same about my (A)-(C). After all, my (B) and (C) are equivalent to his (2) and (3). Admittedly, (A) is weaker than his (1). So I should explain why I need only my (A), and not his (1), for the consequentializing project to succeed.

I argue that any remotely plausible nonconsequentialist theory can be consequentialized. Nevertheless, my consequentializing project could succeed even if this were, as Tenenbaum's (1) states, false. For all I need for my project to succeed is for it to result in a consequentialist theory that has acceptable implications and isn't a mere notational variant of some nonconsequentialist theory. After all, if that's the result, then the project has yielded exactly what I'm looking for: "a moral theory that comports with our considered moral convictions without having to forfeit, as nonconsequentialist theories do, the compelling idea that an act's deontic status is determined by the agent's reasons for preferring its outcome to those of the available alternatives such that it can never be morally wrong for her to act so as to bring about the outcome that she has most reason to want to obtain" (8). So, from here on, I'll focus on my (A)-(C), rather than Tenenbaum's (1)-(3), and look to see if it constitutes a trilemma for me.

One theory that fails in way (B) is *Kantsequentialism*:

S's ϕ -ing is morally permissible if and only if, and because, there is no alternative ψ such that w_{ψ} outranks w_{ϕ} , where saying that w_{ψ} outranks w_{ϕ} is just to say that S's ϕ -ing would be morally impermissible in any situation in which S's ψ -ing is an option and where S's ϕ -ing would be morally impermissible in a given situation if and only if, and because, S's ϕ -ing in that situation violates Kant's Categorical Imperative Test.

Kantsequentialism is a theory that is deontically equivalent to Kantianism, but it is only superficially a consequentialist theory. For it is merely a notional variant of Kantianism, a nonconsequentialist theory. It doesn't say anything different or beyond what Kantianism says; it just says the same thing using different words. The lesson here is that if the consequentializer were to define a consequentialist theory as one that holds that S's ϕ -ing is morally permissible if and only if, and because, there is no alternative ψ such that w_{ψ} outranks w_{ϕ} and then claim that to say that w_{ψ} outranks w_{ϕ} is just to say that S's ϕ -ing would be morally impermissible in any situation in which S's ψ -ing is an option, then the consequentialist project would fail. But I avoid falling into this trap, for I deny that saying that w_{ψ} outranks w_{ϕ} is merely to say that S's ϕ -ing would be morally

impermissible in any situation in which S 's ψ -ing is an option. Instead, I hold that to say that w_ψ outranks w_ϕ is to say that S has more reason to desire w_ψ than to desire w_ϕ , which is not at all equivalent to saying that S 's ϕ -ing would be morally impermissible in any situation in which S 's ψ -ing is an option. Since I understand ' w_ψ outranks w_ϕ ' in this way, I avoid being impaled on horn (B).

Tenenbaum concedes as much, but he worries that I'll now be impaled on horn (C). He says, "'more reason to desire that an outcome obtains (or that a possible world is actualized)' is not clearly an ordinary notion and thus we might suspect that our understanding of this notion is simply derivative of our understanding of a reason to act, or as Portmore puts it 'more reason to perform (an action)'." The worry, then, is that we have no grasp of whether or not S has more reason to desire w_ψ than to desire w_ϕ independent of our thinking that S 's ϕ -ing would be morally impermissible in any situation in which S 's ψ -ing is an alternative. And if that's right, then the consequentializer is doomed to be impaled on horn (C).

But it seems that we do have intuitions about what it is appropriate to desire that's independent of our intuitions about what it is appropriate to do. For instance, it seems that I ought to prefer the possible world in which Pablo saves my child to the possible world in which Pablo saves Otto's two children even though it seems that Pablo ought, other things being equal, to save Otto's two children instead of my child. So which possible world I ought to prefer is not simply a function of what the given agent (in this case, Pablo) ought to do. Perhaps, though, Tenenbaum would claim that it is a function of what I ought to do in some hypothetical world. That is, Tenenbaum could claim that if it were possible for me to cause Pablo to save my child rather than Otto's two children, I ought to do so and that this is what explains why I ought to prefer the possible world in which I cause Pablo to save my child. But it's not clear how this is supposed to explain why I ought to prefer the possible world in which Pablo decides on his own to save my child, which is a different possible world than the one in which I cause Pablo to save my child. Moreover, there are plenty of cases where our intuitions about what state of affairs an agent ought to desire can't possibly be a function of our intuitions about whether that agent ought to act so as to bring that state of affairs about. For instance, it seems that I ought to desire that my wife remembers our anniversary without my causing her to remember it. This intuition couldn't possibly derive from the intuition that I ought to cause her to remember our anniversary without my causing her to remember it, for such isn't even logically possible.¹¹

At this point, Tenenbaum would claim that insofar as we can get a grip on what agents ought to prefer that's independent of our antecedent views about what agents ought to do, these intuitions have unacceptable implications when plugged into BC, thereby resulting in the consequentialist project being impaled horn (A). For instance, Tenenbaum thinks that I ought to prefer my child's getting the job to the most qualified

¹¹ I thank Jamie Dreier for suggesting this example. For another case in which our intuitions about what an agent ought to desire couldn't possibly derive from our intuitions about what she or any other agent ought to do, see *Commonsense Consequentialism*, notes 11 and 75 on pp. 60 and 114, respectively.

applicant's getting the job even though I ought to give the job to the most qualified applicant. Now, it is not clear what particular case he has in mind. Of course, there are clear cases in which I ought to give the job to the most qualified applicant. Certainly, if lives are at stake (if, for instance, I'm hiring a neurosurgeon), then I should give the job to the most qualified applicant even if that's not my child. But, in this case, it seems to me that I ought also to prefer the world in which my child is unemployed and the most qualified applicant gets the job to the world in which my child is employed and lives are put at risk. In other cases, it's clear that I ought to prefer that my child gets the job. If, for instance, I'm hiring a butcher and my child is a recovering drug addict who badly needs a job to get her life back on track, then it seems that I ought to prefer that my child gets the job. But it also seems to me that I ought to give it to her over the most qualified applicant, especially if the other applicant is going to have a much easier time finding alternative employment given her less sullied past. Nevertheless, I must admit that there may be cases where our intuitions about what agents ought to prefer will yield counterintuitive implications when plugged into BC.¹² But this is where the difference between Tenenbaum's (1) and my (A) is crucial. The consequentialist theory that results from the process described above needn't deliver the same deontic verdicts that we started with. Rather, it need only deliver deontic verdicts that are on reflection acceptable—or, at least, more acceptable than denying either MR or TCR.

It may be that the consequentialist project will be impaled on horn (A): that is, it may be that in the end the process described above will result in a consequentialist theory that delivers unacceptable verdicts. I just don't think that Tenenbaum has given us sufficient reason to think that if the project is to avoid both horns (B) and (C), this must be the case.¹³

¹² See, for instance, *Commonsense Consequentialism*, note 75 on p. 114.

¹³ I'm grateful to Daniel Star for chairing the APA session from which this symposium spawned and for suggesting to *Philosophy and Phenomenological Research* that they do a symposium on my book. I thank Brie Gertler for organizing this symposium. And I thank Dale Dorsey, Josh Gert, Paul Hurley, Sergio Tenenbaum, Daniel Star, Jamie Dreier, and Howard Nye for helpful discussions.