

Consequentialism and Moral Rationalism

DOUGLAS W. PORTMORE

IN THIS PAPER, I make a presumptive case for *moral rationalism*: the view that agents can be morally required to do only what they have decisive reason to do, all things considered.¹ And I argue that this view leads us to reject all traditional versions of act-consequentialism. I begin by explaining how moral rationalism leads us to reject utilitarianism.

§1 The too-demanding objection: How moral rationalism leads us to reject utilitarianism

Utilitarianism holds that an act is morally permissible if and only if it maximizes aggregate utility.² This view is too demanding. It implies that agents are morally required to sacrifice their projects, interests, and special relationships whenever doing so would produce more, even just slightly more, aggregate utility than not doing so would. Thus, according to utilitarianism, I'm morally required to sacrifice my life, to neglect my relationship with my daughter, and to abandon my project of completing this paper if I could thereby produce more, even just slightly more, aggregate utility. To demand that I make such sacrifices for the sake of such miniscule gains in aggregate utility is to demand more from me than can be rightfully or reasonably demanded of me.

To say that a given theory is *too* demanding is not merely, or even necessarily, to say that it demands quite a lot from agents in certain circumstances. After all, almost all moral theories demand quite a lot from agents in at least some circumstances.³ What's more, a theory can be too demanding in part because some of its demands, though quite small, are

¹ I assume, contrary to COPP 1997, that there is a normative standpoint from which we can judge what an agent has decisive reason to do, all things considered—in other words, that there is a normative standpoint from which we can judge what an agent just plain ought to do. See MCLEOD 2001 for a reply to Copp.

² The aggregate utility produced by an act is the sum of all the utility it produces minus the sum of all the disutility it produces, where utility is a measure of whatever it is that enhances a subject's welfare, and disutility is a measure of whatever it is that diminishes a subject's welfare. An act maximizes aggregate utility just when there is no available alternative act that would produce more aggregate utility than it would. And note that I use 'utilitarianism' as shorthand for 'maximizing act-utilitarianism'.

³ Paul Hurley (2006, p. 681) makes this point as well.

more than can be rightfully or reasonably demanded of agents. Such is the case with utilitarianism.

To illustrate, imagine that we live in a utopian world in which everyone is not only materially well-off but also extremely happy. And assume, as is in fact the case, that in addition to two other computers I own an ultra-portable laptop computer that I use only for travel. This computer is admittedly a luxury, which I use to pass the time on planes and in airports. To demand that I forfeit this luxury for the sake of promoting the greater good is not to impose any great demand on me. Yet a theory that required me to hand over my computer to some even more well-to-do person who already has *four* computers, including one ultra-portable laptop, would be too demanding. It's one thing to think that, in a non-utopian world, I would be required to give up such a luxury for the sake of saving some poor child's life, but it's quite another to think that, in a utopian world, I would be required to give up such a luxury for the sake of providing some well-to-do person with even more luxury than I have. Nevertheless, utilitarianism requires me to do just that if that's what would maximize utility, as where, for instance, this well-to-do person would get slightly more enjoyment out of my laptop than I would. And this is true even if the net gain to be had is but one measly *utile*, and let me just stipulate that a *utile* is the smallest possible unit of utility: equivalent to someone's experiencing the mildest of pleasures for the briefest of moments. To demand that I make such a sacrifice for the sake of benefiting someone who is even better off than me is to demand more from me than can rightfully or reasonably be demanded of me.

The foregoing suggests that there are two possible senses in which utilitarianism might be too demanding: (1) it holds that agents are morally required to make sacrifices that they are not, in fact, morally required to make or (2) it holds that agents are morally required to make sacrifices that they do not have decisive reason to make, all things considered.⁴ For now, let's call these *sense₁* and *sense₂*, respectively. The claim that utilitarianism is too demanding in *sense₁* is, perhaps, the more common of the two, but it is certainly not the only sense in which utilitarianism is claimed to be too demanding. To see this, consider that some utilitarians concede that their theory is too demanding but deny that this constitutes an objection to it.⁵ If

⁴ In thinking about utilitarianism's being too demanding in *sense₂*, I have been greatly influenced by DORSEY 2009.

⁵ See, for instance, SIDGWICK 1966, SINGER 1999 (pp. 289, & 308-9), and SOBEL 2007 (p. 14). To be fair, I should note that Sobel never officially endorses consequentialism, let alone utilitarianism specifically. Nevertheless, we can glean from what he says that he thinks both

their claim was that utilitarianism is too demanding in sense₁, then this would make no sense, for if utilitarianism holds that we are morally required to make sacrifices that we are not, in fact, morally required to make, then it would indeed be false. One can, however, admit that utilitarianism is too demanding in sense₂ and yet deny that the theory is false. The thought would be that morality is itself too demanding, and so even if utilitarianism is too demanding, it could, nonetheless, be the correct moral theory.

Samuel Scheffler makes the point quite eloquently:

Suppose that morality as represented by a certain philosophical theory strikes us as too demanding in what it requires of people. Does this, in principle, provide the basis for a legitimate objection, either to the theory or to morality itself? If so, is the problem that the theory has distorted the content of morality? Or is it that morality itself is excessively demanding, so that while the theory may be an accurate representation of the content of morality, people have reason to treat moral considerations as less weighty or authoritative than we may previously have supposed? (1992, p. 5)

But what does it mean to say that morality is itself too demanding? Here, the claim would seem to be that morality is too demanding in sense₂. As David Sobel puts it, “What morality asks, we could say, is too much to be the thing the agent has most reason to do all things considered, but not too much to count as what morality asks” (2007, p. 14). So, clearly, there are these two senses in which a moral theory might be thought to be too demanding. Let us call those theories that are too demanding in sense₁ *erroneously demanding* and call those theories that are too demanding in sense₂ *unreasonably demanding*.

I want to focus exclusively on the idea that utilitarianism is unreasonably demanding, for this is something that most philosophers (utilitarians and non-utilitarians alike) can agree on.⁶ Given any plausible conception of practical reasons, utilitarianism is unreasonably demanding, for agents often do not have decisive reason to make the sorts of sacrifices that utilitarianism requires them to make. For instance, even if killing

that utilitarianism is too demanding in sense₂ and that this is no objection to the view. Interestingly, Paul Hurley has argued that Peter Singer is, despite his explicit denials of moral rationalism, implicitly committed to the view (or at least something very close to it)—see HURLEY 2009.

⁶ Utilitarians that would agree with this include Henry Sidgwick (1966) and Peter Singer (1999, pp. 289, & 308-9). Non-utilitarians that would agree with this include Paul Hurley (2006) and Sarah Stroud (1998).

myself would maximize aggregate utility, it's unlikely that I would have decisive reason to do so if the net gain in aggregate utility would be only one utile. Of course, as I've noted above, one can admit that utilitarianism is unreasonably demanding and deny that this constitutes an objection to the theory. But some do take the fact that utilitarianism is unreasonably demanding as an objection to the theory.⁷

What, then, accounts for the disagreement? It seems to be this: unlike those who deny that a theory's being unreasonably demanding constitutes an objection to that theory, those who think that it does constitute an objection accept the following thesis, which I call *moral rationalism*:⁸

MR If a subject, *S*, is morally required to perform an act, *x*, then *S* has decisive reason to perform *x*, all things considered.⁹

If, on the one hand, MR is true, then the correct moral theory cannot be unreasonably demanding. And this explains why those who accept MR reject utilitarianism on account of its being unreasonably demanding, for MR and the fact that utilitarianism is unreasonably demanding implies that

⁷ For instance, Paul Hurley has, in a recent paper, argued against traditional forms of consequentialism, and these arguments proceed "within the widely held framing conviction that any plausible theory of moral standards must be such that rational agents typically have decisive reasons to avoid what these standards identify as cases of wrongdoing" (2006, p. 704). Sarah Stroud argues against consequentialism on the grounds that it conflicts with the following thesis: "If *S* is morally required to ϕ , then *S* has most reason to ϕ "—see STROUD 1998 (esp. p. 171 and pp. 182-184).

⁸ The thesis that I call *moral rationalism* sometimes goes by other names. David Brink calls MR "the supremacy thesis" (1997, p. 255), Stephen Darwall calls it "supremacy" (2006b, p. 286), Samuel Scheffler calls it "the claim of overridingness" (1992, pp. 52-54), John Skorupski calls it "the principle of moral categoricity" (1999, p. 170), Sarah Stroud calls it the "overridingness thesis" (1998, p. 171), and R. Jay Wallace calls it the "optimality thesis" (2006, p. 130).

⁹ The essential idea is that agents can be morally required to do only what they are rationally required to do. Now, some philosophers (see, e.g., STROUD 1998) formulate this thesis using 'most reason' in place of 'decisive reason'. This formulation is equivalent to mine unless either some reasons have no (rational) requiring strength (see GERT 2004 and DANCY 2004) or some reasons with requiring strength are silenced, undermined, or bracketed off by other factors or considerations (see, e.g., SCANLON 1998). If either is the case, then *S*'s having most reason to perform *x* would not entail *S*'s being rationally required to perform *x*. And, in that case, I assume that Stroud and the others would want to replace 'most reason' with 'decisive reason', where '*S* has decisive reason to perform *x*' just means '*S*'s reasons are such as to make *S* (objectively) rationally required to perform *x*'. In other words, *S* has decisive reason to perform *x* if and only if *S* lacks sufficient reason to perform any alternative to *x*.

the theory is erroneously demanding and, thus, false. But if, on the other hand, MR is false, then the correct moral theory can be unreasonably demanding without being erroneously demanding. And this explains why those who reject MR are content with the fact that utilitarianism is unreasonably demanding.

So far, we've seen how the observation that utilitarianism is unreasonably demanding will lead those who accept MR to reject utilitarianism. Of course, I haven't yet provided any reason to think that MR is true. In §3, I'll rectify this by making a presumptive case for MR. But before proceeding with that, I want to show how the objection that utilitarianism is unreasonably demanding is really just a species of a much more general objection.

§2 The argument against utilitarianism from moral rationalism

The idea that utilitarianism is objectionable insofar as it is unreasonably demanding is really just a species of a much more general objection, for the objection proceeds on the assumption that MR is true, and once we accept MR, we should object to utilitarianism whenever it requires agents to perform acts that they don't have decisive reason to perform, whether these acts be self-sacrificing acts or even self-benefiting acts. Thus, we should object to utilitarianism not only on the grounds that it sometimes requires agents to make sacrifices that they don't have decisive reason to make, but also on the grounds that it sometimes requires agents to perform self-benefiting acts that they don't have decisive reason to perform.

To illustrate, let's suppose that a woman named Anita would benefit from murdering her estranged uncle. Assume that she's crafty enough to get away with it and that, if she did, she would finally receive the heart transplant that she so desperately needs, for her uncle is an organ donor with the same rare blood group and tissue type that she has. Let's further suppose that murdering her uncle would maximize aggregate utility, for let's assume that the gain in utility resulting from Anita's heart transplant would slightly more than offset the loss in utility resulting from her uncle's murder. Given this further stipulation, utilitarianism requires Anita to benefit herself by murdering her uncle. Arguably, though, she does not, in this instance, have decisive reason to commit murder.

So, if we accept MR, we should object to utilitarianism because it sometimes requires agents to perform self-benefiting acts that they don't have decisive reason to perform. We can, then, offer the following general

schema for constructing an argument against utilitarianism on the basis of MR:

- C1 There exists an act, x_1 , available to S that would maximize aggregate utility but that S does not have decisive reason to perform, all things considered.
- C2 If utilitarianism is true, then S is morally required to perform any act that would maximize aggregate utility. (From the definition of 'utilitarianism')
- C3 So, if utilitarianism is true, then S is morally required to perform x_1 . (From C1 and C2)
- C4 S is morally required to perform x_1 only if S has decisive reason to perform x_1 , all things considered. (From MR)
- C5 So, S is not morally required to perform x_1 . (From C1 and C4)
- C6 Therefore, utilitarianism is not true. (From C3 and C5)

The argument is deductively valid, and the only controversial premises are C1 and C4. Since C4 follows from MR and since I'm interested in what follows if MR is true, I'll focus on C1. To see just how plausible C1 is, consider the following two plausible candidates for x_1 .

First, let x_1 be the act of your sacrificing the life of your partner (whose life, we'll assume, is well worth living) for the sake of saving some stranger's life, where your doing so would maximize aggregate utility, but where the net gain in aggregate utility would be quite small: one utile. Suppose, for instance, that both lives contain nearly the same amount of utility but that what accounts for there being more utility in saving the stranger's life is the fact that her life contains slightly more utility: an extra second of mild pleasure.

Clearly, in this case, you lack decisive reason to save the stranger's life, for the reason that you have to save your partner is at least as strong as, if not stronger than, the reason you have to bring about an extra second of mild pleasure. After all, let's assume that you have a very special relationship with your partner, involving years of personal history together and mutual love and respect for one another. Assume (1) that you care more about the welfare of your partner than you do about the welfare of this stranger, (2) that you care more about the welfare of your partner than you do about producing small net gains in aggregate utility, and (3) that saving your partner is what would best serve your own interests. On what plausible theory of practical reasons, then, would there not be at least sufficient reason to save your partner instead of this stranger? I can't think

of any. Thus, this seems to be a clear case where a subject, S (viz., you), could maximize aggregate utility by performing an act, x_1 (specifically, the act of saving the stranger instead of your partner), but where x_1 is not something S has decisive reason to do.

Second, let x_1 be the act of murdering your partner so as to ensure that each of a billion and one people get some minor benefit. Assume that, if you murder your partner, the loss in utility would be one billion utiles, and assume that the benefit to each of the billion and one people would be one utile. Thus, the net gain to be had in your murdering your partner is but one utile. To make C1 even more plausible, assume that your partner is one of society's worst off and undeservedly so and that the billion and one people who you could instead benefit are society's most well-to-do and undeservedly so. In fact, let's assume that they are well-to-do only as a result of exploiting society's worst off, which includes your partner.

The idea that, in this situation, all the various reasons that you have to refrain from murdering your partner are, even when taken together, insufficient to successfully counter the fairly weak reason that you have to produce a net gain of one utile is quite implausible. Consider all the reasons that you have to refrain from murdering your partner: (1) the reason that you have to respect your partner's autonomy, (2) the reason that you have to give priority both to your own welfare and to the welfare of those whom you love, (3) the reason that you have to give priority to society's worst off, and (4) the reason that you have to give priority to the deserving over the undeserving. Given how much reason you have not to murder your partner and how little reason you have to produce a small net gain in aggregate utility, we should deny that you have decisive reason to murder your partner.

Given these two plausible candidates for x_1 , we can be confident that there are at least two acts that would have both the property of being what would maximize aggregate utility and the property of not being what the agent has decisive reason to do. So we should accept C1, and, therefore, acknowledge that utilitarianism cannot be the correct moral theory if MR is true. The problem for utilitarianism vis-à-vis MR is that utilitarianism ignores two important classes of rather weighty reasons—reasons that, in certain instances, have sufficient weight to decisively oppose the reason that one has to maximize utility: (1) reasons that have nothing to do with *promoting* utility, such as the reason one has to ensure that utility is fairly distributed and (2) reasons that stem from the special relations that we bear to ourselves and our loved ones, such as the reason that we have to favor ourselves and our loved ones when deciding whose utility to promote. If a

theory of morality, such as utilitarianism, ignores such reasons, it will conflict with MR.¹⁰

And it's not just utilitarianism that ignores whole classes of reasons. All forms of traditional act-consequentialism do so as well. According to all forms of *traditional act-consequentialism* (TAC), we should accept the following *criterion of rightness* (cr):

TAC_{cr} An act is morally permissible if and only if it maximizes the good (impersonally construed).

TAC_{cr} ignores two important classes of rather weighty reasons—reasons that, in certain instances, have sufficient weight to decisively oppose the reason that one has to maximize the good: (1) reasons that have nothing to do with promoting the good, such as the reason one has to refrain from violating someone's autonomy even when doing so is a means to promoting the good, and (2) reasons that stem from the special relations that we bear to ourselves and our loved ones, such as the reason one has to promote the good by saving one's own loved one as opposed to by helping some stranger save her loved one.¹¹

Insofar as we think either that we have reasons to do things besides promote the good or that the weight of our reasons to perform various good-promoting acts depends on our relationship to those whose good would thereby be promoted, TAC_{cr} fails to acknowledge certain reasons and/or their proper weights.¹² And this means that TAC_{cr} will conflict with MR. Indeed, we could easily revise the argument that I gave against utilitarianism above so as to constitute an argument against traditional act-consequentialism. We need only substitute 'traditional act-consequentialism' for 'utilitarianism' and 'the good' for 'aggregate utility' in the above argument. I leave this as an exercise for the reader.

§3 The presumptive case for moral rationalism

¹⁰ This sort of argument comes from STROUD 1998.

¹¹ Even if the traditional act-consequentialist holds, say, that there is more value in a person's life being saved by a loved one than there is in a person's life being saved by some stranger, she must deny that there is, other things being equal, any gain in value to be had by saving one's own loved as opposed to helping someone else to save her loved one.

¹² I borrow this point from STROUD 1998. Paul Hurley (2009) makes a similar point, although he focuses on the fact that TAC_{cr} ignores "non-impersonal reasons." I think, though, that the challenge is broader than that, for it seems that a person could consistently hold that there are impersonal reasons that have nothing to do with promoting the good.

So far, we've seen how moral rationalism (MR) compels us to accept consequentialism while at the same time leading us to reject all traditional versions of the theory. Of course, these observations are of little interest unless MR is true, and I have yet to provide any reason for thinking that it is. Below, I attempt to rectify this. Let me forewarn the reader, though, that the case that I'll be making in this paper is, at best, a presumptive one.¹³ My arguments will establish only that those who reject MR must also reject an intuitively plausible premise. Thus, there is a price to be paid in rejecting MR. Of course, there is often a price to be paid in rejecting a philosophical thesis. And, perhaps, some will think that, in this instance, the price is worth paying. Myself, I don't think that it is.

Before I can make the presumptive case for MR, though, I'll need to explain some of the relevant terminology. First, to say that S performs $\sim x$ is to say that S performs the "act" of refraining from performing x . So if x is, for instance, the act of pushing some specific button at some particular time, then $\sim x$ is the "act" of refraining from pushing that button at that time.

Second, to say that S has *decisive reason* to φ is to say that S's reasons are such as to make S objectively rationally required to φ , and to say that S has *sufficient reason* to φ is to say that S's reasons to φ are such as to make S objectively rationally permitted to φ . Thus, S has decisive reason to φ if and only if S lacks sufficient reason to $\sim\varphi$. I'll use 'S has decisive reason to φ ' and 'S has sufficient reason to φ ' as shorthand for 'S has decisive reason to φ , all things considered' and 'S has sufficient reason to φ , all things considered', respectively.

Third, to say that S has sufficient reason to perform $\sim x$ is to say that there is at least one act-token that is an instance of S's refraining from performing x that S has sufficient reason to perform. Likewise, to say that S has decisive reason to perform $\sim x$ is to say that there is at least one act-token that is an instance of S's refraining from performing x that S has decisive reason to perform.

¹³ One reason that the case that I'll be making is, at best, a presumptive one is that I don't, for reasons of space, attempt to address all the various objections and counterarguments that have been leveled against MR. For instance, Shaun Nichols (2002) has argued that a recent empirical study, which he himself conducted and which probes people's intuitions about psychopaths, undermines MR. Although I don't have space here to address Nichols's interesting argument, I think that others have adequately defended MR against his argument. See, for instance, KENNETT 2006. For more on this issue, see also JOYCE 2008a, NICHOLS 2008, and JOYCE 2008b.

Fourth, to say that S *freely performs* x (or $\sim x$)¹⁴ is to say that S performs x —having the relevant sort of control over whether or not she performs x —that is, the sort of control that is necessary for her being an appropriate candidate for praise or blame with respect to her having performed x . More specifically, to say that S freely performs x is to say that S satisfies whatever other conditions in addition to having the relevant sort of knowledge (such as knowledge about what she is doing and about what she is bringing about) that is necessary for her to be morally responsible for having performed x .

Fifth, to say that S *knowledgeably performs* x is to say that S performs x knowing all the relevant facts—the relevant facts being those facts the ignorance of which would either inculpate or exculpate her for performing x .

Sixth, to say that S is *blameworthy* for performing x is to say both that it is appropriate for S to feel guilt about having performed x and that it is appropriate for others to feel indignation—and, perhaps, also resentment—in response to S's having performed x . And, here, I use 'appropriate' in the sense of being apt, fitting, or correct and, thus, in the same sense that fear is the appropriate response to the perception of danger. In this sense, it can be appropriate to blame oneself or someone else even though having this attitude (and/or expressing it) would not be instrumental in bringing about any good.¹⁵

Seventh and last, when I say that S is morally required to perform x , I mean this in the objective sense—that is, I mean that S is *objectively* morally required to perform x .

With these terms defined, I can now state the argument as follows:

- C7** If S is morally required to perform x , then S would be blameworthy for freely and knowledgeably performing $\sim x$.¹⁶
- C8** S would be blameworthy for freely and knowledgeably φ -ing only if S does not have sufficient reason to φ .¹⁷

¹⁴ Hereafter, I'll leave the '(or $\sim x$)' implicit.

¹⁵ It's important to keep distinct the issue of whether it is appropriate to blame someone for doing x and the issue of whether it is appropriate to intend to perform an act that constitutes the outward expression of this attitude. One can be appropriate without the other being appropriate. See PORTMORE 2010 (§§3.2.3) for more on this matter. And see BENNETT 1979, STRAWSON 1962, and WALLACE 1994 for criticisms of the view that blame is appropriate if and only if the outward expression of this attitude would lead to a desired change in the agent and/or her behavior.

¹⁶ 'S would be blameworthy for freely and knowledgeably performing $\sim x$ ' should be read as shorthand for 'S would be blameworthy if S were to freely and knowledgeably perform $\sim x$ '.

- C9** So, if S is morally required to perform x , then S does not have sufficient reason to perform $\sim x$. (From C7 and C8)
- C10** If S does not have sufficient reason to perform $\sim x$, then S has decisive reason to perform x . (From the definitions of ‘sufficient reason’ and ‘decisive reason’ above)
- C11** Therefore, if S is morally required to perform x , then S has decisive reason to perform x —and this is just MR. (From C9 and C10)¹⁸

The argument is deductively valid, and it seems that C8 is the only assumption that isn’t a conceptual truth. C10 is clearly a conceptual truth. And C7 expresses the common assumption that there is a conceptual connection between blameworthiness and wrongdoing.¹⁹ Although there may not be an essential connection between blameworthiness and wrongdoing *per se*, there is, it seems, an essential connection between blameworthiness and *freely and knowledgeably* doing what’s wrong. And if this is right, then anyone who denies MR is committed to denying C8. Yet C8 is intuitively plausible—even those who deny MR admit as much.²⁰ Besides, there is, I think, a sound argument for C8, which is based on assumptions that are even more intuitively plausible than C8 is. But before I present that argument, let me rebuff some putative counterexamples to C8.²¹

¹⁷ Note that C8 entails: ‘S would be blameworthy for freely and knowledgeably performing $\sim x$ only if S does not have sufficient reason to perform $\sim x$, all things considered’. One need only substitute $\sim x$ for φ .

¹⁸ This sort of argument for MR is not original to me. Darwall (2006a and 2006b, p. 292) and Skorupski (1999, pp. 170-171) make similar arguments for MR. And Shafer-Landau (2003, pp. 192-193) presents the same kind of argument but for a considerably weaker thesis: viz., if S is morally required to perform x , then S has a (not necessarily decisive or even sufficient) reason to perform x .

¹⁹ See, for instance, DARWALL 2006a, GIBBARD 1990, MILL 1991 [1861], and SKORUPSKI 1999.

²⁰ For instance, David Sobel, who explicitly denies MR in SOBEL 2007, claims in SOBEL 2009 that it “seems quite intuitive that earnestly blaming a person for φ -ing entails the view that the agent all things considered ought not to have φ -ed.” And he says, “It also seem quite intuitive to say that if one is acting as one has most reason to act, then one is acting as one ought and therefore one’s action is not worthy of blame.”

²¹ Jussi Suikkanen and Paul McNamara have both suggested that C8 might have problematic implications in the following sort of case. Suppose that I accidentally (and, perhaps, also non-negligently) make two incompatible appointments: A and B. They’re incompatible, because it’s impossible for me to keep both them given that they’re for the same time and in different locations. Assume, though, that A and B are equally important. Given their equal importance, it may seem that I have sufficient reason to break A as well as sufficient reason to break B. And, thus, according to C8, neither would be blameworthy.

Consider the following example. Suppose that Arthur, a white supremacist, sneaks up behind an unsuspecting black man, named Bert, and clubs him over the head, knocking him unconscious. He does so out of hatred for blacks. However, unbeknownst to Arthur, Bert was just about to shoot and kill his ex-girlfriend Carla, who, we'll suppose, is completely innocent. As it turns out, then, Arthur's act saves Carla's life. Arthur had, then, sufficient reason to club Bert over the head, for assume that any less violent or injurious act would have been insufficient to save Carla's life. Nevertheless, Arthur is clearly blameworthy. So this may seem to be a counterexample to C8. But note that Arthur doesn't *knowledgeably* club Bert over the head. Arthur is ignorant of the fact that his clubbing Bert over the head is necessary to save Carla's life, and it is his ignorance of this fact that inculpates him. The antecedent in C8 is, therefore, false, and so this is no counterexample to C8.

Of course, a better counterexample is certainly in the neighborhood. Imagine, for instance, a slightly different version of the case. In this version, Arthur knows all the relevant facts and, thus, freely *and knowledgeably* clubs Bert over the head. Nevertheless, what motivates him is not the thought that doing so will save Carla but only the desire to hurt a black person. Here, too, it seems that Arthur had sufficient reason to club Bert over the head and yet is clearly blameworthy.²² But although it's clear

Yet, surely, this is a moral dilemma if ever there was one, and so I would be blameworthy no matter which appointment I break. So C8 seems to rule out moral dilemmas. But, in fact, as Howard Nye has pointed out to me, C8 rules out moral dilemmas only if there are no rational dilemmas, and it's unclear why anyone would want to claim that there can be moral dilemmas but no rational dilemmas. So if someone wants to claim that I would be blameworthy no matter which appointment I break, then that someone should also claim that I lack sufficient reason to break either appointment and so would be doing something contrary reason no matter which appointment I break.

²² I thank Mike Almeida, Richard Chappell, and Clayton Littlejohn for proposing these sorts of putative counterexamples. Another type of putative counterexample, suggested to me by Peter de Marneffe, is Sophie's choice. It may seem both that Sophie had sufficient reason to choose to save her son and that it is, nevertheless, appropriate for her to feel guilty for having made this choice. This, however, is a tricky case. My suspicion is that although we may think her guilt appropriate insofar as we question whether she truly had sufficient reason to sacrifice her daughter, we shouldn't think it *appropriate* for her to feel guilt (in the sense of its being apt as opposed to psychologically normal) insofar as we think that she did have sufficient reason to make the choice that she made. But it does seem reasonable to question whether she in fact had sufficient reason to make the choice that she made. We might think, for instance, that she had decisive reason to give each child an equal chance at being saved and that it is, therefore, appropriate for her to feel guilty insofar as she may have failed to have done so—perhaps, she chose to save her son because he was her

that Arthur is blameworthy, it's far from clear that Arthur is blameworthy for clubbing Bert over the head. I think that we can rightly blame Arthur for his vicious motive, for his malevolent intent, and for his racist attitudes.²³ We can even rightly blame him for acting out of hatred and malice. But I don't think that we can rightly blame him for clubbing Bert over the head, as this is exactly what he should have done. Later in life, when Arthur finally comes to realize the error of his ways, what he should come to regret and feel guilty about is the fact that he was a racist who acted out of hatred and malice. Arthur should not, however, regret having clubbed Bert over the head, nor should he feel guilty for having done so, for this is what he had sufficient (indeed, decisive) reason to do. To accept blame for having φ -ed, one must judge that one should not have φ -ed. But although he should neither have wished Bert harm nor acted out of hatred or malice, he should have clubbed him over the head, as this was necessary to save Carla's life.

In arguing that Arthur should be blamed, not for clubbing Bert over his head, but for his malicious intent, his racist attitudes, and his acting out of malice, I've appealed to C8. This may seem illegitimate, but I'm trying to establish only that if one finds C8 intuitively plausible, the above examples shouldn't dissuade one from accepting C8. Although it's clear that Arthur is blameworthy for something, it's not clear that he is blameworthy for freely and knowledgeably doing anything that he had sufficient reason to do. Thus, the proponent of C8 can reasonably claim that there is no φ such that Arthur had sufficient reason to φ and yet is blameworthy for freely and knowledgeably φ -ing. If, on the one hand, we let ' φ ' stand for 'clubbing Bert over the head', then although it's clear that Arthur had sufficient reason to φ , it's not so clear that he is blameworthy for freely and knowledgeably φ -ing. And if, on the other hand, we let ' φ ' stand for something such as 'acting out of malice', then although it's clear that Arthur is blameworthy for freely and knowledgeably φ -ing, it's not so clear that he had sufficient reason to φ .²⁴ So I don't see any clear

favorite. Or we might think that a mother never has sufficient reason to sacrifice her child, not even for the sake of ensuring that at least one of her two children lives.

²³ If necessary, we should assume that Arthur is morally responsible for these mental acts/states either because they are directly under his control or because they are the result of actions (or inactions) that were directly under his control—e.g., the result of his decision not to perform certain acts that were essential to the proper development of his good moral character. (As to why this might be unnecessary, see note 26.)

²⁴ Note that saying that Arthur acted out of malice is not equivalent to saying that Arthur clubbed Bert over the head. These are not two equivalent descriptions of the same act, for Arthur can club Bert over the head without acting out of malice.

counterexample to C8. Having cleared that up, let me explain why we should accept C8.

The thought underlying C8 is that agents are blameworthy only for freely and knowledgeably doing what they lack sufficient reason to do. One way to bring out the intuitive plausibility of this claim is to point to the tension there is in blaming someone for acting in a certain way while acknowledging that she had sufficient reason to act in that way. Stephen Darwall puts the point thusly:

It seems incoherent...to blame while allowing that the wrong action, although recommended against by some reasons, was nonetheless the sensible thing to do, all things considered.... Part of what one does in blaming is simply to say that the person shouldn't have done what he did, other reasons to the contrary notwithstanding. After all, if someone can show that he had good and sufficient reasons for acting as he did, it would seem that he *has* accounted for himself and defeated any claim that he is to blame for anything. Accepting blame involves an acknowledgment of this proposition also. To feel guilt is, in part, to feel that one shouldn't have done what one did. (2006b, p. 292)²⁵

Another way to bring out the intuitive plausibility of C8 is to point to the tension there is in holding someone morally responsible for her actions on account of her having the capacity to respond appropriately to the relevant reasons and then blaming her for responding appropriately to the relevant reasons by doing what she had sufficient reason to do. Let me explain. It seems that an agent can be blameworthy for her actions only if she is morally responsible for them and that she can be morally responsible for them only if she has the relevant sort of control over her actions.²⁶

²⁵ See also DARWALL 2006a, especially p. 98.

²⁶ Some deny that control is a necessary condition for moral responsibility. Some argue that we can appropriately be held responsible for our judgment-sensitive attitudes even if we do not have volitional control over them—see, for instance, SCANLON 1998, SMITH 2005, and SMITH 2008. The thought is that an agent is morally responsible for something if and only if that something is rightly attributable to her in that it expresses her judgments, values, or normative commitments. This view is known as attributionism, and it contrasts with volitionalism (see, e.g., LEVY 2005). But even attributionists allow that volitional control can be relevant in determining moral responsibility for bodily movements. Angel M. Smith, for instance, says: “In some cases (e.g., in determining a person’s responsibility for a bodily movement), it may make sense to ask whether the agent has *voluntarily chosen* the thing in question, because that will determine whether that thing can reasonably be taken to express her judgments” (2008, p. 368). In any case, what matters to attributionists is whether the action is connected to the agent’s underlying normative judgments such that she can, in principle, be called upon to defend it with reasons and be blamed if no adequate defense can be provided (SMITH 2008, p. 370). Given that citing sufficient reason for performing an

What's more, it's plausible to suppose that she has the relevant sort of control over her actions only if she has the capacity to respond appropriately to the relevant reasons for action—that is, the capacity both to recognize what the relevant reasons are and to react appropriately to them, being moved by them in accordance with their associated strengths.²⁷ Indeed, it's this capacity for responding appropriately to the relevant reasons, which normal adult humans typically possess and which children, animals, and the criminally insane typically lack, that distinguishes those who can potentially be blameworthy for their actions from those who can't.

Now, given that an agent can be blameworthy only if she has the capacity to respond appropriately to the relevant reasons, it's important to note that, in flawlessly exercising this capacity, she could be led to perform any act that she has sufficient reason to perform. Indeed, if there is but one act that she has sufficient reason to perform, then this capacity will, if exercised flawlessly, lead her both to recognize that this is the only act that she has sufficient reason to perform and to react by performing it. But if an agent is morally responsible and, thus, potentially blameworthy in virtue of having the capacity to respond appropriately to the relevant reasons, then how can we rightly blame her for doing something that she has sufficient reason to do when, in flawlessly exercising this very capacity, this is what she is led to do?

To put the point in slightly different terms, it seems inappropriate to hold an agent responsible on the condition that she has the capacity to be guided by sound practical reasoning and then blame her for acting as she might very well be led to act if she is guided by sound practical reasoning. And since an agent can be led to perform any act that she has sufficient reason to perform when guided by sound practical reasoning, it seems inappropriate to blame her for freely and knowledgeably doing what she has sufficient reason to do.

More formally, the argument for C8 is this:

action is presumably an adequate defense of one's performing that action, attributionists seem committed to my claim that S would be blameworthy for freely and knowledgeably φ -ing only if S does not have sufficient reason to φ —that is, to C8.

²⁷ We may need to talk about the capacities of mechanisms rather than the capacities of agents—see FISCHER AND RAVIZZA 1998, especially p. 38. If so, my argument could easily be revised so as to replace all talk of the agent's capacity to respond appropriately to the relevant reasons with talk of the reasons-responsiveness of the mechanism that issued in the agent's action.

- C12** S is morally responsible for whether or not she ϕ s and, thus, potentially blameworthy for ϕ -ing or failing to ϕ only if S has the relevant sort of control over whether or not she ϕ s.
- C13** S has the relevant sort of control over whether or not she ϕ s only if S has the capacity to respond appropriately to the relevant reasons (both moral and non-moral).²⁸
- C14** So, S is potentially blameworthy for whether or not she ϕ s only if S has the capacity to respond appropriately to the relevant reasons. (From C12 and C13)
- C15** If S has sufficient reason to ϕ , then, in flawlessly exercising her capacity to respond appropriately to the relevant reasons, S could be led to freely and knowledgeably ϕ .²⁹
- C16** If S is potentially blameworthy for whether or not she ϕ s only if she has the capacity to respond appropriately to the relevant reasons, then S can't be blameworthy for freely and knowledgeably ϕ -ing when, in flawlessly exercising this capacity, S could be led to freely and knowledgeably ϕ .
- C17** So, S wouldn't be blameworthy for freely and knowledgeably ϕ -ing if S has sufficient reason to ϕ . (From C14, C15, and C16³⁰)
- C18** Therefore, S would be blameworthy for freely and knowledgeably ϕ -ing only if S does not have sufficient reason to ϕ —and this is just C8. (From C17 by contraposition)

In considering this argument, it will be helpful to have a specific example in mind. So consider the case in which I've promised to meet with a student to discuss his exam grade but have the unique opportunity to make some sum of money by instead giving a lecture. Assume that I cannot both give the lecture and keep my promise to the student. Assume that I'm offered just enough money so that I have just as much reason, all

²⁸ For a further defense of C12 and C13, see FISCHER AND RAVIZZA 1998. Note that it's not enough for an agent to have the capacity to respond appropriately to non-moral reasons. To be morally responsible, an agent must have the capacity to respond, and respond appropriately, to moral reasons as well. Certain psychopaths have the capacity to respond appropriately to non-moral reasons but are incapable of recognizing the fact that the rights and interests of others provide them with reasons for acting in ways that respect their rights and promote their interests. Such psychopaths are not morally responsible for their actions—see FISCHER AND RAVIZZA 1998, pp. 76-81.

²⁹ Whether she would or not depends on whether she also has sufficient reason to ϕ . It could be, after all, that she has sufficient reason both to ϕ and to refrain from ϕ -ing.

³⁰ Roughly, the structure of the argument from C14-C16 to C17 goes something like this: $B \rightarrow C, S \rightarrow L, (B \rightarrow C) \rightarrow (L \rightarrow \sim B) \therefore S \rightarrow \sim B$.

things considered, to give the lecture as to keep my promise to meet with the student. And assume that I have considerably less reason to do anything else. That is, assume that I have both sufficient reason to keep my promise and sufficient reason to give the lecture and insufficient reason to do anything else. Furthermore, assume that I'm strongly responsive to reasons, meaning both that I always recognize what the relevant reasons are and that I'm always moved to act in accordance with their associated strengths, such that I always do what I have decisive reason to do, only do what I have sufficient reason to do, and choose arbitrarily which of two acts to perform (as by, say, the toss of a coin) if and only if I have sufficient and equivalent reason to do either. Thus, being strongly reasons-responsive and having sufficient and equivalent reason to do either, I choose arbitrarily to give the lecture. That is, I first designate heads for giving the lecture and designate tails for keeping my promise and then toss a coin, which lands heads. As a result, I decide to give the lecture.

Of course, being strongly reasons-responsive I would have kept my promise and met with the student had there been decisive reason to do so. For instance, I would have kept my promise had my moral reason for doing so been a bit stronger, as where, say, the student's scholarship was at stake. Likewise, I would have kept my promise had my self-interested reason for giving the lecture been a bit weaker, as where, say, the monetary offer had been a bit less. And had there been a third option that I had decisive reason to do (e.g., saving some drowning child), I would have done that instead. But, as it was, none of these were the case. As it was, I had sufficient reason to break my promise and give the lecture. And that's what I did, having flawlessly exercised my capacity for sound practical reasoning.

Given my flawless execution of my capacity for sound practical reasoning, how can I be faulted for breaking my promise? Isn't it inappropriate to hold me morally responsible and, thus, potentially blameworthy in virtue of my capacity for being guided by sound practical reasoning and then blame me for acting as sound practical reasoning leads me to act? Shouldn't even the student admit that, had he been in my situation and perfectly rational, he might also have been led to act as I did? And if so, how can he rightfully resent me for acting as he would have acted?

Of course, things would have been different had there been decisive reason for me to have kept my promise. In that case, it would have been appropriate for the student to resent me for breaking my promise to him. If I had decisive reason to keep my promise as well as the capacity to

respond appropriately by recognizing this fact and reacting appropriately to it, then I must have failed to have done something that I could, and should, have done. Perhaps, I failed to recognize the fact that I had decisive reason to keep my promise even though I was capable of doing so. Or, perhaps, I recognized that I had decisive reason to do so, but failed to react appropriately even though I had the capacity to do so—perhaps, I was weak-willed and gave into the temptation to earn some easy money. In either case, I would have failed in some way and could, then, rightfully be blamed for this failure. So I see how it can be appropriate to blame someone for failing to do what she had decisive reason to do, but I don't see how we can rightfully blame someone for freely and knowledgeably doing what she had sufficient reason to do.

Admittedly, lots of people perform acts that they have sufficient reason to perform and are, nonetheless, blameworthy, because they were motivated by some consideration that shouldn't have motivated them. Such is the case in the example from above where Arthur does what he has sufficient reason to do out of a desire to see people of a certain race harmed. But, as I noted above, I think that in this case we should not blame Arthur for doing anything that he had sufficient reason to do (such as, clubbing Bert over the head). The alternative would be to allow that people can be blameworthy for freely and knowledgeably doing what they had sufficient reason to do, and that, I've argued, is inappropriate insofar as the very capacity that makes them subject to blame can lead them, when exercised flawlessly, to do what they have sufficient reason to do. The only way to forestall such a possibility is to insist upon MR, which, I've argued, we should.

Still, the person who wishes to deny MR might balk at C13, arguing that when it comes to being morally responsible, what's relevant is not the capacity to respond in a *rationally* appropriate manner, but rather only the capacity to respond in a *morally* appropriate manner.³¹ Of course, this person must insist that the two can come apart and that, although doing what one recognizes as one's having sufficient reason to do is always a rationally appropriate response, it is sometimes a morally inappropriate response.³² But why think this? Why think that it is morally inappropriate

³¹ To be clear, 'to respond appropriately' should be taken as shorthand for 'to respond in a rationally appropriate manner' both in C13 and throughout the rest of the argument.

³² I thank Pat Greenspan for raising this possible objection to C13. As far as I know, however, no one has defended such a position in print. Even those who think that being morally (or normatively) competent is a necessary condition for being morally responsible stop well short of suggesting that this competency requires the capacity to respond by

to respond to moral reasons as it is rationally appropriate to respond to them? A moral reason is just a certain type of reason, and, as a reason, it has a certain weight. So why think that it is morally inappropriate to give moral reasons only as much weight as they in fact have?

In any case, when we talk of an agent's needing control over her actions in order to be held morally responsible for them, the relevant sort of control seems to be *rational* self-control—that is, the capacity to respond in a rationally appropriate manner to the relevant reasons. If it were anything else, then someone could be strongly reasons-responsive and, *for that reason*, lack the sort of control that's essential for moral responsibility. But the idea that an agent could fail to be morally responsible on account of her being strongly reasons-responsive is quite counterintuitive.³³

*School of Historical, Philosophical, & Religious Studies
Arizona State University
P.O. Box 874302
Tempe, AZ 85287-4302*

giving moral reasons *more* than their due rational weight. Rather, they suggest, to the contrary, that agents must have the ability to grasp and respond to both moral *and* non-moral reasons—see, for instance, FISCHER AND RAVIZZA 1998 (chap. 3) and WOLF 1990 (chap. 6). Indeed, at least one author—viz., Susan Wolf—explicitly claims that responsibility entails the ability to appreciate and act in accordance with *all* the reasons there are, including both moral and non-moral reasons—see WOLF 1990, p. 141. On Wolf's Reason View of responsibility, “full freedom and responsibility will involve the ability to appreciate reasons that come from a variety of sources” (1990, p. 141) and she explicitly notes that “our image of the agent who is most able to see and understand what reasons there are need not coincide with that of the agent who is most acutely sensitive particularly to *moral* reasons” (1990, p. 137). And although she claims that freedom and responsibility involves the ability to appreciate and act in accordance with what she calls the “True and the Good,” she makes clear that “[a]ppreciation of the Good need not be confined to appreciation of the *moral* good. Indeed, in certain contexts, appreciation of the moral good may interfere with one's ability to appreciate the nonmoral good or with one's ability to recognize reasons for preferring a morally inferior course of action” (1990, p. 137). Thus, being overly sensitive to moral reasons can diminish one's freedom and responsibility by diminishing one's rational self-control.

³³ For helpful feedback on earlier versions of this paper, I thank Richard Chappell, Peter de Marneffe, Thomas Hurka, Paul McNamara, G. Shyam Nair, Robert Neal, Howard Nye, Ben Sachs, Andrew Schroeder, Mark Schroeder, Russ Shafer-Landau, Dave Shoemaker, David Sobel, Jussi Suikkanen, Peter Railton, Toni Rønnow-Rasmussen, Travis Timmerman, Mark van Roojen, and audiences at both the 2010 Arizona Workshop for Normative Ethics and the 2010 New Orleans Invitational Seminar in Ethics.

douglas.portmore@asu.edu

References

- BENNETT, J. (1979). "Accountability." In Z. V. Straaten (ed.), *Philosophical Subjects: Essays Presented to P. F. Strawson*, pp. 14-47. Oxford: Oxford University Press.
- BRINK, D. O. (1997). "Kantian Rationalism: Inescapability, Authority, and Supremacy." In G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason*, pp. 255-291. New York: Oxford University Press.
- COPP, D. (1997). "The Ring of Gyges: Overridingness and the Unity of Reason." *Social Philosophy and Policy* 14: 86-106
- DANCY, J. (2004). "Enticing Reasons." In R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, pp. 91-118. Oxford: Oxford University Press.
- DARWALL, S. (2006a). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- DARWALL, S. (2006b). "Morality and Practical Reason: A Kantian Approach." In D. Copp (ed.), *The Oxford Handbook of Ethical Theory*, pp. 282-320. Oxford: Oxford University Press.
- DORSEY, D. (2009). "Weak Anti-Rationalism and the Demands of Morality." Forthcoming in *Noûs*.
- FISCHER, J. M. AND RAVIZZA, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- GERT, J. (2004). *Brute Rationality: Normativity and Human Action*. Cambridge: Cambridge University Press.
- GIBBARD, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.
- HURLEY, P. E. (2009). *Beyond Consequentialism*. Oxford: Oxford University Press.
- HURLEY, P. E. (2006). "Does Consequentialism Make Too Many Demands, or None at All?" *Ethics* 116: 680-706.

- JOYCE, R. (2008a). "What Neuroscience Can (and Cannot) Contribute to Metaethics." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality*, pp. 371-394. Cambridge, MA: MIT Press.
- JOYCE, R. (2008b). "Responses to Nichols and Katz." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality*, pp. 419-426. Cambridge, MA: MIT Press.
- KENNETT J. (2006). "Do Psychopaths Really Threaten Moral Rationalism?" *Philosophical Explorations* 9: 69-82.
- LEVY, N. (2005). "The Good, the Bad and the Blameworthy." *Journal of Ethics & Social Philosophy*, www.jesp.org, vol. 1, no. 2.
- MCLEOD, O. (2001). "Just Plain 'Ought'." *The Journal of Ethics* 5: 269-291.
- MILL, J. S. (1991). [1861]. *Utilitarianism*. In J. M. Robson (ed.), *Collected Works of John Stuart Mill*, vol. 10, pp. 203-59. London: Routledge.
- NICHOLS, S. (2008). "Moral Rationalism and Empirical Immunity." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality*, pp. 395-408. Cambridge, MA: MIT Press.
- NICHOLS, S. (2002). "How Psychopaths Threaten Moral Rationalism, Or Is It Irrational to be Amoral?" *The Monist* 85: 285-304.
- PORTMORE, D. W. (2010). *Commonsense Consequentialism: Wherein Morality Meets Rationality*. Unpublished manuscript available at <http://www.public.asu.edu/~dportmor/Commonsense%20Consequentialism.htm>.
- SCANLON, T. M. (1998). *What We Owe to Each Other*. Cambridge, Mass.: The Belknap Press.
- SCHEFFLER, S. (1992). *Human Morality*. New York: Oxford University Press.
- SHAFFER-LANDAU, R. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.
- SIDGWICK, H. (1966). [1907]. *The Methods of Ethics*. New York: Dover.
- SINGER, P. (1999). "A Response." In D. Jamieson (ed.), *Singer and His Critics*, pp. 269-335. Oxford: Blackwell.
- SKORUPSKI, J. (1999). *Ethical Explorations*. Oxford: Oxford University Press.
- SMITH, A. M. (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115: 236-271.

- SMITH, A. M. (2008). "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 138: 367-392.
- SOBEL, D. (2009). "Subjectivism and Blame." Forthcoming in the *Canadian Journal of Philosophy*.
- SOBEL, D. (2007). "The Impotence of the Demandingness Objection." *Philosophers' Imprint* 7: 1-17.
- STRAWSON, P. (1962). "Freedom and Resentment." *Proceedings of the British Academy* 48: 187-211.
- STROUD, S. (1998). "Moral Overridingness and Moral Theory." *Pacific Philosophical Quarterly* 79: 170-189.
- WALLACE, R. J. (2006). *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason*. Oxford: Oxford University Press.
- WALLACE, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press.
- WOLF, S. (1990). *Freedom within Reason*. New York: Oxford University Press.