

Maximalism and Rational Control*

Douglas W. Portmore

Abstract: Maximalism is the view that if an agent is permitted to perform a certain type of action (say, baking), this is in virtue of the fact that she is permitted to perform some instance of this type (say, baking a pie), where φ -ing is an instance of ψ -ing if and only if φ -ing entails ψ -ing but not vice versa. Now, the point of this paper is not to defend maximalism, but to defend a certain account of our options that when combined with maximalism results in a theory that both avoids the sorts of objections that have typically been levelled against maximalism and accommodates the plausible idea that a moral theory must be collectively successful in the sense that everyone's satisfying the theory guarantees that our theory-given aims will be best achieved. I argue that, for something to count as an option for an agent, it must, in the relevant sense, be under her control. And I argue that the relevant sort of control is the sort that we exercise over our reasons-responsive attitudes (e.g., our beliefs, desires, and intentions) by being both receptive and reactive to reasons. I call this sort of control *rational control*, and I call the view that φ -ing is an option for an agent if and only if she has rational control over whether she φ s *rationalism*. When we combine this view with maximalism, we get *rationalist maximalism*, which I argue is a promising moral theory.

Keywords: collectively successful, maximalism, morality, obligations, options, principle of moral harmony, prisoner's dilemma, Professor Procrastinate, rationality, rational control, reasons-responsiveness, Regan, voluntary control.

The performance of one option can entail the performance of another. That is, the performance of the one can logically necessitate the performance of the other. For instance, I have both the option of baking and the option of baking a pie, and baking a pie entails baking, for it is logically impossible to bake a pie without baking. Such instances of performance entailment are ubiquitous. Kissing passionately entails kissing. Walking while chewing gum entails walking. Driving under 55 mph entails driving under 100 mph. Marrying an unmarried man entails marrying a bachelor. Stretching a t_1 and then going for a run at t_2 entails going for a run at t_2 .

Given that our options include both those that entail others and those that are entailed by others, our theories of morality and rationality owe us an account of the permissibility of each type. Moreover, they owe us an account of which, if either, is more fundamental than the other: the permissibility of the option that entails the other (e.g., baking a pie) or the permissibility of the option that it entails (e.g., baking)? Suppose, for instance, that it is permissible both to bake and to bake a pie. Is baking permissible because it is permissible to bake a pie, or is baking a pie permissible because it is permissible to

* Draft of October 13, 2015. The latest draft can always be found at <http://bit.ly/1NmZ57E>.

bake? Or are they equally fundamental, as they would be if they were both permissible because they were both, say, in accord with Kant's categorical imperative?

On one view, the permissibility of an option that entails another is always more fundamental than the permissibility of the option that it entails. This view is sometimes called *maximalism*.¹ On this view, we must distinguish between maximal and non-maximal options. An option φ is a *maximal option* if and only if there is no option ψ such that ψ -ing entails φ -ing but not vice versa. Otherwise, it's a *non-maximal option*. So walking at t_1 won't be a maximal option if walking fast at t_1 is an option. And walking fast at t_1 won't be a maximal option if walking fast at t_1 and then jogging at t_2 is an option. It seems, though, that we will eventually arrive at an option that isn't entailed by any other option. That's a maximal option.

We must distinguish between these two types of options, because, on maximalism, they are to be evaluated differently.

Maximalism: (Max1) For any non-maximal option ν , S 's ν -ing is permissible if and only if there exists an option φ such that S 's φ -ing is permissible and S 's φ -ing entails S 's ν -ing but not vice versa, and when S 's ν -ing is permissible, this is in virtue of the fact that there exists an option φ such that S 's φ -ing is permissible and S 's φ -ing entails S 's ν -ing but not vice versa. And, (Max2) for any maximal option μ , S 's μ -ing is permissible if and only if S 's μ -ing has feature F , and when S 's μ -ing is permissible, this is in virtue of the fact that S 's μ -ing has feature F .

Maximalism is neutral on what sorts of things count as options. Perhaps, it is only voluntary acts. But, perhaps, the non-voluntary formations of beliefs and other attitudes also count as options. And maximalism allows for options to be conjunctive. An agent could, for instance, have the option of asserting that p while believing that p is false. What's more, maximalism is neutral on what 'F' stands for. Indeed, we can substitute for 'has feature F' anything that would render Max2 coherent, including 'maximizes utility', 'accords with

¹ See, for instance, Bykvist 2002 and Gustafsson 2014. Note, however, that my use of the term is broader than theirs. Whereas they use the term to refer only to consequentialist theories that take the permissibility of maximal options to be more fundamental than that of non-maximal options, I use the term to refer to any theory that takes the permissibility of maximal options to be more fundamental than that of non-maximal options.

Kant's categorical imperative', or 'contains only beliefs for which S has sufficient evidence and contains all and only those acts that maximize expected utility'. As this last example illustrates, F can be something quite complicated, involving the assessment of different types of options according to different criteria—e.g., evaluating beliefs in terms of evidence and acts in terms of expected utility.

The thought underlying maximalism is that if I'm permitted to perform a certain type of action (say, baking), then I must be permitted to perform some instance of this type (say, baking a pie). For if I'm not permitted to bake anything—not a pie, not a cake, not cookies, not anything, then I'm not permitted to bake. Moreover, if I'm not permitted to bake a pie and then eat either some or none of it (and assume that, if I bake a pie, I must then eat either some or none of it), I'm not permitted to bake a pie. Suppose, for instance, that someone will kill me and everyone I love if either I bake a pie and then eat some of it or bake a pie and then eat none of it. In that case, it would be impermissible for me to bake a pie. And this is because there is no instance of baking a pie that I'm permitted to perform. Note, then, that, as I'm understanding things, S's φ -ing counts as an instance of S's ψ -ing if and only if S's φ -ing entails S's ψ -ing but not vice versa. Thus, both baking-a-pie-and-then-eating-some-of-it and baking-a-pie-and-then-eating-none-of-it count as instances of baking a pie.

So, on maximalism, if I'm permitted to bake, this is because I'm permitted to perform some instance of baking, such as baking a pie. And if I'm permitted to bake a pie, this is because I'm permitted to perform some instance of pie-baking, such as baking an apple pie. And if I'm permitted to bake an apple pie, this is because I'm permitted to perform some instance of apple-pie-baking, such as baking an apple pie and then taking it to the family who just moved in across the street. And so on and so forth. But, of course, if this just goes on forever, we'll end up with an infinite regress. I believe, however, that this will not go on forever. We will eventually arrive at an option that is not entailed by any other option.

Some may question my belief. The worry would be that, for any option (e.g., the option of punching), there will always be a more specific option that entails it (e.g., the option of punching a punching bag) and an even more specific option that entails that option (e.g., the option of punching a punching bag softly), and so on and so forth, *ad*

infinitum. It may seem, then, that we will never arrive at an option that is so specific that there is no other more specific option that entails it.

But although there is no limit to the degree of specificity with which we can describe a particular action, there is a limit to an agent's ability to determine the specificity of her actions. Consider that although I can determine whether I punch a punching bag with my right or left fist and also whether I punch it softly or forcefully, I can't, it seems, determine whether I punch it with precisely 100.0002 newtons of force. And, importantly, it seems that I must be able to control whether I ϕ if my ϕ -ing is to count as an option for me and, thus, as something that I can be obligated to perform and be accountable for failing to perform. But, in the relevant sense, I do not control whether I punch the bag with precisely 100.0002 newtons of force, and so this does not count as an option for me. And this in turn explains why I cannot be obligated to do so and why I can't be held accountable should I fail to do so. So even though I have the option of punching the bag softly and my punching the bag with 100.0002 newtons of force entails punching the bag softly, I don't have the option of punching it with precisely 100.0002 newtons of force. For the control that I exert over the specificity of my actions is limited. Consequently, the specificity of my options is limited. Thus, it is not the case that, for any option that I have, I must have some other more specific option that entails it. Although there will always be a more specific action that entails it, that more specific action will not always be an option. And this along with the fact that our existences are finite suggests that we will always eventually arrive at a maximal option.

When we do arrive at a maximal option, we won't be able to derive its permissibility from that of some other permissible option that entails it. Given that it's a maximal option, there will be no other option that entails it. Thus, maximalism must include not only Max1 but also Max2. Max2 tells us that a maximal option is permissible if and only if it has some feature F. Moreover, Max2 tells us that when a maximal option is permissible, this is in virtue of the fact that it has feature F. But what's distinctive and interesting about maximalism is that it tells us that we don't evaluate non-maximal options in the same way. Instead of evaluating non-maximal options in terms of whether or not they have feature F, we evaluate them in terms of whether there is some other permissible option that entails it. Thus, Max1 tells us that a non-maximal option is permissible if and only if there is some permissible option that entails it but not vice versa. Moreover, it says that when a non-

maximal option is permissible, this is in virtue of the fact that there is some permissible option that entails it *but not vice versa*. The need for this italicized phrase can be seen when we consider pairs of options that each entail the other—e.g., marrying a bachelor and marrying an unmarried man. It can't be that the one is permissible in virtue of being entailed by the other. For to say that an option is permissible in virtue of being entailed by itself (or its logical equivalent) is to offer no explanation at all. So what explains the permissibility of marrying an unmarried man is not that it is permissible to marry a bachelor, but that it is, say, permissible to marry some specific bachelor (i.e., some specific unmarried man).

Now, the point of this paper is not to argue for maximalism. That has been done elsewhere.² Instead, the point is to argue for a certain account of what our options are and to show that when we combine this account with maximalism, we arrive at view that can both avoid some typical worries associated with maximalism and accommodate the idea that a moral theory ought to be collectively successful (or morally harmonious)—that is, something that it would be good for everyone to follow. In section 1, I argue that, for something to count as an option for an agent, it must, in the relevant sense, be under her control. And I argue that the relevant sort of control is the sort that we exercise over our reasons-responsive attitudes (e.g., our beliefs, desires, and intentions) by being both receptive and reactive to reasons—forming, revising, sustaining, and/or abandoning these attitudes in light of our awareness of facts (or what we take to be facts) that count for or against them.³ I call this sort of control *rational control*, and I call the view that ϕ -ing is an

² See Feldman 1986, Goldman 1978, Portmore 2013, Portmore 2015, and Zimmerman 1996.

³ The notion of a *reasons-responsive attitude* is, perhaps, the same as Thomas M. Scanlon's notion of a *judgment-sensitive attitude*, an attitude that is sensitive to the subject's judgments about reasons (1998, 20). But Scanlon's notion is, if not distinct, misleading, for we can respond to reasons without having any judgments about what our reasons are. "We respond to reasons when we are aware of facts that give us these reasons, and this awareness leads us to believe, or want, or do what these facts give us reasons to believe, or want, or do" (Parfit 2011, 493). Thus, we can respond to reasons while neither knowing that this is what we are doing nor having any judgments about our reasons (Parfit 2011, 461).

Reasons-responsive attitudes include all and only those mental states that a rational subject will tend to have, or tend not to have, in response to reasons (or apparent reasons)—facts (or what are taken to be facts) that count for or against the attitudes in question. So beliefs are clearly reasons-responsive attitudes, for a rational subject will, for instance, tend to believe that it will rain in response to her awareness of facts that constitute decisive reasons for her believing this, such as the fact that a reliable weather service has predicted that it will rain. Although reasons-responsive attitudes include many mental states, they exclude feelings of hunger, nausea, tiredness, and dizziness, which are not responsive to reasons. Suppose, for instance, that I have too quickly consumed a good-sized meal and am still feeling hungry, as there has not yet been sufficient time for my brain to receive the relevant physiological signals from my stomach. Even if I am aware that I've eaten more than enough to be satiated, my hunger is not responsive to this awareness. Instead, it is responsive only to the physiological signals that supposedly take about twenty minutes to travel from the stomach to the brain.

option for a subject if and only if she has rational control over whether she φ s *rationalism*. In section 2, I show that when we combine rationalism with maximalism, we arrive at a view—viz., *rationalist maximalism*—that can avoid the sorts of objections that have typically be levelled against maximalism. And, in section 3, I explain the sense in which we should expect a moral theory to be collectively successful. And I argue that rationalist maximalism is uniquely well-suited to meet this condition. I conclude, therefore, that rationalist maximalism is a promising moral theory.

1. What are our options?

There is some set of mutually exclusive events (or alternatives) such that, if one of its members, viz., φ , is more highly favored by the relevant considerations than any other member, then S ought to φ . I call this the set of S's options. As noted above, some of these options will be maximal and some will be non-maximal. And, if maximalism is correct, these two types of options are to be assessed differently. But before we can even get to assessing them, we need to know what they are. They are, I believe, all and only those things are, in some relevant sense, under our control. This would explain why, for instance, I ought to spend the weekend writing a lecture even though the relevant considerations (e.g., those concerning what would be most beneficial to me and others) favor my spending it writing a literary masterpiece. Unfortunately, given my lack of literary talent, whether I write a literary masterpiece is not under my control. And, thus, it can't be something that I ought to do. It is only those metaphysically possible actions over which I exert control that can be things that I ought to do. And, of those, my spending the weekend writing a lecture is, or so we'll assume, the most highly favored by the relevant considerations. Thus, it is what I ought to do.

The thought, then, is that S's having control at t over whether she φ s at t' is both necessary and sufficient for φ -ing at t' to be, as of t , an option for her, which in turn is necessary for her to be obligated, as of t , to φ at t' and to be accountable should she fail to φ at t' ($t < t'$).⁴ Now, it's important to make explicit, as I have done here, the relevant time indices, for what one has control over—and, thus, what constitutes an option for one—can

⁴ As I think of them, temporally-indexed options and obligations refer to properties that are possessed by the agent at certain times. Thus, the phrase "S is, as of t , obligated to φ " is equivalent to "S has at t the property of being obligated to φ ." And, likewise, the phrase " φ -ing is, as of t , an option for S" is equivalent to "S has at t the property of having φ as an option."

vary over time. For instance, I used to have the option of never setting foot in Europe. And, at the time, I had control over whether I would ever set foot in Europe. But now that I have already set foot in Europe, I no longer have the option of never setting foot in Europe. At this point, I control only whether I set foot in Europe again.

Since I'm claiming that an agent's having control at t over whether she φ s at t' is both necessary and sufficient for φ -ing at t' to be, as of t , an option for her, I need to defend both the idea that it is necessary and the idea that it is sufficient. To see that it is necessary, consider a view that denies this.

Schedulism: A subject S has, as of t , the option of φ -ing at t' if and only if she would φ at t' if she were, leading up to t' , to have certain intentions at certain times. More precisely, for any event φ , subject S , and times t , t' , t'' , and t''' , S has at t the option of φ -ing at t''' if and only if, and because, there is some schedule of intentions I extending over a time-interval T beginning at t' such that the following are all true: (a) if S 's intentions followed I , then S would carry out all the intentions in I ; (b) S 's carrying out all the intentions in I would logically necessitate S 's φ -ing at t''' ; (c) S has, as of t , the capacity to continue, or to come, to have the intentions that I specifies for t' ; and (d) for any time t'' in T after t' , if S 's intentions followed I up until t'' , then S would have just before t'' the capacity to continue, or to come, to have the intentions that I specifies for t'' ($t < t' < t'' < t'''$).⁵

On this view, an agent's having control at t over whether she φ s at t' is not necessary for φ -ing at t' to be, as of t , an option for her. To illustrate, suppose that no matter what I do now and no matter what I intend now to do tomorrow, I'm not going to get up and exercise when the alarm wakes me at 5 AM. Instead, I'm going to hit the snooze button several times, getting up at 6 AM, leaving myself with no time to exercise. Still, on schedulism, I have now the option of getting up and exercising at 5 AM so long as (a) I will at 5 AM have the capacity to form the intention to get up and exercise and (b) would get up and exercise if I were to form that intention. The fact that no matter what I do now or intend now to do later, I won't

⁵ This is adapted from Jacob Ross's notion of a performable option—see his 2012, 81. Proponents of schedulism include Fred Feldman (1986) and Michael J. Zimmerman (1996).

form the intention to get up and exercise when the alarm wakes me is, on schedulism, irrelevant. But this seems a mistake. In order for my getting up and exercising at 5 AM tomorrow to be at present an option for me, I must now have control over whether I will get up and exercise at 5 AM tomorrow. But, if I'm not going to get up and exercise at 5 AM tomorrow no matter what my present actions and intentions are, then I don't now have control over whether I'll get up and exercise at 5 AM tomorrow. After all, to have control now over what I'll do at 5 AM tomorrow is to have the present ability to affect what I'll do then. But, in this case, I have, at present, no way of affecting whether I'll get up and exercise at 5 AM tomorrow, or so we're assuming.

To better understand where schedulism goes wrong, consider the following case.

Curing Cancer: A five-year-old boy named Saru is playing with a computer keyboard. Millions of lives depend on his writing and sending an email to the National Institutes of Health that explains how to cure cancer in such a way that those who read it must take it seriously.⁶

Schedulism implausibly implies that Saru has the option of curing cancer. After all, to cure cancer, he need only write and send this email, which in turn just involves his making a certain series of keystrokes. And Saru would make each keystroke in that series if he were to have certain intentions at certain times. To illustrate, suppose that such an email would start with: "The cure for cancer is...." If at t_1 Saru were to intend to hit Shift + T, he would do so at t_2 , thereby typing an uppercase T. And, having done that, he would have the capacity to form at t_3 the intention to hit the H key. Moreover, if he were to form this intention at t_3 , he would then hit the H key at t_4 , thereby typing a lowercase H. And similar assumptions apply for all the remaining keystrokes in the series. Thus, schedulism implies that Saru has, as of t_0 , the option of curing cancer. And since the relevant considerations favor Saru's doing so, it seems that the proponent of schedulism must insist that Saru is obligated to cure cancer and can be held accountable should he fail to do so. This, of course, is absurd. Saru cannot control whether he types out the cure for cancer. For, as we'll plausibly assume, Saru will not make the required series of keystrokes no matter what he

⁶ This example is modeled after a similar one given in Wiland 2005.

does now or intends now to do later. Even if he does happen to hit Shift + T at t_2 , he wouldn't (or so I'll assume) follow up by making each of the other ten thousand or so specific keystrokes that are required to complete the task of writing and sending such an email. And even if Saru were to intend to type out (or to try to type out) the cure for cancer, he would fail (or so I'll assume). And although God could perhaps form the complex intention to make the many specific keystrokes required, Saru does not have the capacity to form such a complex intention. Thus, Saru has no way at t_0 to effect the typing out of such an email. No present intention or action that he has the capacity to form or perform would result in his curing cancer. Thus, he does not now control whether he cures cancer. And the problem with schedulism is that it holds that Saru has, at present, the option of curing cancer even though he does not now control whether he does so. It seems, then, that we should hold that an agent's having control at t over whether she φ s at t' is necessary for φ -ing at t' to be, as of t , an option for her.

And, as noted above, having control seems sufficient as well. To see this, consider a view that denies this.

Tryism: A subject S has, as of t , the option of φ -ing at t' if and only if she would φ at t' if she were to try, intend, or decide at t to φ at t' .

To see both that tryism implies that an agent's having control at t over whether she φ s at t' is insufficient for φ -ing at t' to be, as of t , an option for her and that this view is problematic for this very reason, consider the following case.

Stupid Mistake: A genius named Albert took a math test and missed one of the easiest problems because he overthought things and, consequently, overlooked its simple solution.⁷

According to tryism, Albert didn't have the option of providing the correct answer to this easy problem. After all, it wasn't for a lack of trying that he didn't provide the correct

⁷ This is inspired by John Maier's example of a golfer who misses an easy putt, which he uses to question whether it's being the case that S would φ if she were to try to φ is necessary for S's having the option to φ . See his 2014.

answer. Rather, he tried to provide the correct answer and failed because he overthought things, consequently overlooking the most obvious solution. So, if anything, he tried too hard. And given that he was going to try and fail, it was false that “he would provide the correct answer if he were to try do so.” So tryism implies that providing the correct answer wasn’t even an option for him.

But that’s strange. Even Albert would admit that he could and should have got the correct answer to such an easy problem. And we should too. For it seems that he had both the ability and opportunity to provide the correct answer. In other words, it seems that he had control over whether he was to provide the correct answer. The fact that he failed to exercise this control to good effect doesn’t mean that he lacked control. It just means that he messed up. Indeed, these sorts of stupid mistakes cause us the most frustration precisely because we think that we could and should have got things right. When we get a test back and find that the solution was one that would never have occurred to us, we are not nearly as frustrated (if at all) as when we find that the solution was so obvious that it should have occurred to us. This suggests, then, that an agent’s having control at t over whether she φ s at t' is sufficient for φ -ing at t' to be, as of t , an option for her and that we should reject views, such as tryism, that deny this.

Of course, ultimately how plausible my claim that an agent’s having control at t over whether she φ s at t' is both necessary and sufficient for φ -ing at t' to be, as of t , an option for her will depend on what I take the relevant sort of control to be. So I’ll turn now to that issue. As I see it, there are two main contenders: (1) voluntary control and (2) rational control. I’ll explain each in turn.

Voluntary control is the sort of control that we exert directly over our voluntary actions (such as raising one’s arm) and indirectly over those things that we manipulate via such actions (such as the movement of the car that one’s driving). To better understand this kind of control, it’s important to note that voluntary actions have three key features. For if φ is something that I can directly and voluntarily do, it follows that: (1) I can φ at will, and, thus, I can φ simply by trying, deciding, intending, choosing, or otherwise willing to φ ; (2) I can φ for any reason that I take to be sufficient for doing so, and, thus, I can φ to win a bet, to help others, or to please my partner; and (3) I can choose when to φ , and, thus, can choose

to φ now, five minutes from now, a day from now, or only on Tuesdays.⁸ Taking these features into account, I offer the following tentative account of voluntary control: S has at t direct voluntary control over whether she φ s at t' if and only if S has, as of t , the ability to intentionally φ at t' (as well as the ability to intentionally refrain from φ -ing t') for any reason that she takes as counting sufficiently in favor of her doing so. I offer this account, not as something that I'm committed to in its details, but only as an approximation that will be sufficient for our purposes. Thus, voluntary control just is whatever sort of control that we exercise over our voluntary actions and those things that we manipulate via such actions, and this holds regardless of whether I have gotten the details exactly right.

Even with only this rough and ready account to work with, it's clear enough that we don't typically exert voluntary control over our reasons-responsive attitudes. For the sort of control that I have, say, over whether I believe that Aristotle went for a swim on his 30th birthday is clearly distinct from the sort of control that I have, say, over whether I touch my nose. I can do the latter but not the former at will. Thus, I can do the latter but not the former to win a bet. And whereas I can choose when to touch my nose, I cannot choose when (or even whether) to believe that Aristotle went for a swim on his 30th birthday.⁹

The fact that we don't typically exert voluntary control over our reasons-responsive attitudes along with the plausible thought that people ought to form certain reasons-responsive attitudes and can be held accountable if they fail to do so has led several philosophers to conclude that the sort of control that's relevant to determining our obligations and responsibilities is not voluntary control but rational control.¹⁰ By 'rational control', I just mean whatever sort of control that we exert directly over our reasons-responsive attitudes and indirectly over those things that we influence via such attitudes (such as our voluntary actions). We exercise control over our reasons-responsive attitudes by being both receptive and reactive to reasons—forming, revising, sustaining, and/or abandoning these attitudes in light of our awareness of facts (or what we take to be facts) that count for or against them.

⁸ See McHugh 2012, McHugh 2014, and McHugh Forthcoming.

⁹ For a further defense of the claim that we don't typically exert voluntary control over our reasons-responsive attitudes as well as a response to those who have argued to the contrary, see McHugh 2012, McHugh 2014, and McHugh Forthcoming.

¹⁰ See Graham 2012 (13, note 22), Hieronymi 2006, McHugh Forthcoming, Scanlon 1998, Smith 2005, and Smith 2015. Whereas Angela M. Smith and I use the term 'rational control', Pamela Hieronymi uses the term 'evaluative control' and Conor McHugh uses the term 'attitudinal control'. The basic idea, though, is the same.

Some may hesitate to call this a kind of control given that control must be active and forming an attitude in response to reasons seems less than fully active. But even if forming an attitude in response to reasons is not as active as, say, performing a voluntary action, it does involve thinking, and thinking is active. Compare, then, the formation of an attitude with the feeling of a sensation (such as hunger or dizziness). The feeling of a sensation is completely passive. We simply suffer our sensations. And, because of this, we can't be asked to justify them. We can be asked only to explain them. By contrast, the formation of an attitude is active. For we shape our attitudes by attending to, reflecting upon, and responding to our reasons. And it is because our attitudes reflect the extent to which we have attended to, reflected upon, and responded to our reasons that we can be asked to justify them — to give our reasons for them. So even if the formation of an attitude in response to reasons isn't active in the same way that a voluntary act is, it is, nevertheless, active. So, we do exert a kind of control over our reasons-responsive attitudes: it's just that it is rational control as opposed to voluntary control.¹¹

But what does this sort of control amount to? Here's a tentative account: S has at t rational control over whether she ϕ s at t'' if and only if she has, as of t , the capacity to respond appropriately at t' to the relevant reasons and whether she ϕ s at t'' depends (in the right way) on whether and how she responds at t' to these reasons ($t \leq t' < t''$).¹² Here, too, I offer this only as a rough approximation and not as something that I committed to in all its details. Rational control just is whatever sort of control we exert by attending to, reflecting upon, and responding to reasons, regardless of whether I've got the specifics right.

Given these two types of control, we have the following two competing accounts of what our options are.

Voluntarism: S's having at t voluntary control over whether she ϕ s at t' is both necessary and sufficient for ϕ -ing at t' to be at t an option for her ($t < t'$).

¹¹ The ideas in this paragraph are taken from Hieronymi 2006. See also Hieronymi 2008 and McHugh Forthcoming.

¹² As formulated, this seems to presume that the relevant sort of control is regulative control rather than guidance control, the difference being that you can have the latter, but not the former, with respect to ϕ even when not- ϕ -ing is not an option. But, again, I'm not committed to this or to any other detail concerning this account. If we have only guidance control, and not regulative control, with respect to our reasons-responsive attitudes, then rational control should be thought of as a kind of guidance control. For more on the distinction between these two types of control and its relevance, see Fischer & Ravizza 1998. For how we might think of rational control as a kind of guidance control, see McHugh 2013 and McHugh 2014.

Rationalism: S's having at t rational control over whether she φ s at t' is both necessary and sufficient for φ -ing at t' to be at t an option for her ($t < t'$).

Of these two views (and I see no plausible third view), I believe that rationalism is the clear winner. There are at least three reasons for this. First, voluntarism forces us to deny that people ought ever to form, revise, sustain, or abandon their reasons-responsive attitudes. But ordinarily we think that people often ought to do so. We think, for instance, that parents ought to want what's best for their children, that agents ought to intend to take what they believe to be the necessary means to their ends, and that those who believe that the Earth is no more than a few thousand years old ought to abandon their belief in light of the overwhelming scientific evidence to the contrary. Now, since wanting, intending, and believing are reasons-responsive attitudes that are under our rational control, the rationalist can accept such commonsense normative judgments. But, since these attitudes are not under our voluntary control, the voluntarist cannot. The voluntarist must deny, for instance, that those who believe that the Earth is no more than a few thousand years old ought to abandon their belief. For the voluntarist denies that they have the option of abandoning their belief. The best the voluntarist can do, then, is to claim that these people ought to perform whatever voluntary acts would cause them to abandon their belief. But, in many instances, they can't even claim this. For, in many instances, agents ought not to perform the acts that would cause them to abandon such a belief. To illustrate, suppose that, given Jane's penchant for conspiracy theories, the only thing that would cause her to abandon her belief that the Earth is no more than a few thousand years old is to read a book by some quack claiming that the Bible was written and propagated by the CIA for the purposes of controlling the masses. And let's assume that Jane ought not to read this book both because it contains a lot of dangerous misinformation that she is liable to believe and because she promised her mother that she would stop feeding her hysteria by reading such books. Here, then, is a case where Jane ought to abandon her belief and the voluntarist can't even claim that she ought to do what will cause her to abandon this belief. The rationalist, by contrast, has no problem accounting for the fact that Jane ought to abandon her belief. For Jane has, we'll presume, rational control over whether she does so. That is, she has, we'll presume, the

capacity to respond appropriately to the reasons she has for abandoning this belief—that is, those stemming from the scientific evidence of which she is aware. And if she were to respond appropriately to these reasons, she would thereby abandon her belief. It's just that, in this case, she culpably fails to respond appropriately to her reasons.

Second, it's not just that voluntarism forces us to reject our commonsense normative judgments about reasons-responsive attitudes. It also forces us to reject some of our commonsense normative judgments about acts. For many of the acts that we ought to perform are *mixed acts*—acts that have both a voluntary component and a non-voluntary component.¹³ Such acts include acting in good faith, offering a sincere apology, and expressing one's gratitude. To perform such acts, we must have certain reasons-responsive attitudes. For instance, we can't act in good faith without having the intention to follow through with our part of the bargain. We can't offer a sincere apology without feeling contrite. And we can't express our gratitude without feeling grateful. So whereas we may have voluntary control over whether we say the words "I'm sorry" or "Thank you," we don't have voluntary control over whether we offer a sincere apology or express our gratitude. For we don't have voluntary control over whether we feel contrite or grateful, which is essential to performing such mixed acts. The voluntarist must, therefore, deny that we have the option of performing such mixed acts and, so, deny both that we ought to perform such acts and that we can appropriately be held accountable when we fail to do so. This seems like a substantial cost for those who accept voluntarism. Fortunately, rationalism doesn't come at such a high price. Since we have rational control, for instance, over both whether we utter the words "I'm sorry" and over whether we feel contrite, the rationalist holds that we have the option of performing such mixed acts and, thus, can be obligated to perform them and responsible for failing to perform them.

Third, voluntarism implausibly implies that manipulated agents, primitive animals, and very young children are to be held responsible for their voluntary actions even if they did not have rational control over the volitions that gave rise to them. For, on voluntarism, what matters in determining our obligations and responsibilities with respect to our actions is just whether we had voluntary control over those actions. To see why this is a mistake,

¹³ Alex King (2014) is, as far as I know, the first to point out that we often presume that we ought to perform such mixed acts even though we cannot perform them at will. I'm indebted to her 2014 for many of the ideas presented in this paragraph.

consider the following case. Suppose that I have both the ability and opportunity to kill my neighbor's dog. And assume that, at present, I have no desire to kill or otherwise harm this dog and that, indeed, I have the same sorts of values with respect to non-human animals that other animal liberationists have. But, now, suppose that an evil neuro-scientist abducts me and performs brain surgery on me, manipulating my brain in such a way that I no longer have the capacity to recognize and respond appropriately to the moral reasons there are for not harming non-human animals. Moreover, he implants in me a device that stimulates my neurons so as to cause me to form a very strong desire to kill dogs. Consequently, when I wake up from the surgery, I kill my neighbor's dog. Now, according to voluntarism, I had the option of refraining from killing this dog. For had I recognized the moral reasons for so refraining and considered them to be sufficient, I would have refrained from killing the dog for these reasons. Indeed, I had the ability to intentionally refrain from killing the dog for any reason that I took to be sufficient. It's just that given the manipulation of my brain I was unable to recognize any moral reason for refraining from killing the dog and so took my desire to kill the dog to be sufficient reason to do so. And since killing the dog was under my voluntary control, I'm accountable for having done so.

But this seems like a case where I didn't have the option of refraining from killing the dog and so can't be held accountable for having done so. After all, I lacked rational control both over whether I formed the desire to kill the dog and over whether I formed the intention to kill the dog. And if I didn't have rational control over these states, then I didn't have rational control over the act that they gave rise to. Fortunately, rationalism captures such intuitions. According to rationalism, I lacked the option of refraining from killing the dog given that I lacked rational control over whether I killed the dog. And rationalism can say similarly plausible things about primitive animals and very young children. They may have voluntary control over their actions, but to the extent that they lack the general capacity to recognize and respond appropriately to reasons or lack the specific capacity to recognize and respond to certain relevant types of reasons (say, moral reasons), they'll bear little to no responsibility for failing to respond appropriately to them.

So a given account of when a subject *S* has the option to φ must get the extension of both variables—'*S*' and ' φ '—right. And the problem with voluntarism is that it gets neither right. First, voluntarism holds that the extension of '*S*' is all and only those subjects who can

do things at will (that is, voluntarily). But it seems that although manipulated agents, primitive animals, and very young children can do things at will, they are not the sorts of subjects that have options in the sense that's relevant to determining their obligations and responsibilities. Rather, it seems that the relevant subjects are, as rationalism would have it, all and only those who have the capacity to recognize and respond appropriately to the relevant sorts of reasons.

Second, voluntarism holds that the extension of 'φ' is all and only those things over which we exert voluntary control—that is, our voluntary acts and those things that we manipulate through our voluntary acts. But our options seem to extend well beyond our voluntary acts and those things that we manipulate through them. It seems, for instance, that we can have obligations and responsibilities with respect to things that are not under our voluntary control, such as whether we form certain reasons-responsive attitudes, whether we perform certain mixed acts, and whether we overlook the most obvious solution to an easy math problem.

At this point, I hope to have convinced the reader that rationalism provides a plausible account of what are options are—one that we should take seriously. I now want to show that when we combine rationalism with maximalism, we end up with a view—viz., rationalist maximalism—that can avoid the standard worries associated with maximalism as well as provide perhaps our best hope for finding a moral theory that is collectively successful.

2. Objections to Maximalism

If we accept rationalism, it's not just voluntary acts that count as options. Mixed acts and the formations of reasons-responsive attitudes also count as options. This, of course, has implications for what our maximal options consist in. My saying "I'm sorry" won't count as a maximal option if I have the option of saying this while feeling contrite. And my saying "I'm sorry" while feeling contrite won't count as a maximal option if I have the option of doing this at t_1 and then promising at t_2 to never do it (the offending act) again. Even this won't count as a maximal option if I have the option of doing all this while intending to never to do it again. And so on, until we reach the limits of my rational control. So our maximal options will consist in complex sets of acts and attitudes. Now, if we want to

combine rationalism with maximalism in a way that accommodates the plausible idea that acts and attitudes are to be assessed using different criteria, we'll need 'F' in Max2 to stand for something rather complicated, such as "includes all and only those attitudes that are fitting and includes all and only those acts that would, if the agent were to have all and only fitting attitudes, maximize the good." The resulting view — viz., rationalist maximalism — will then assess non-maximal options (such as feeling contrite or saying "I'm sorry") in terms of whether there is some permissible maximal option that entails forming or performing these non-maximal options.

Rationalist maximalism is, I believe, the most plausible version of maximalism. For one, rationalism is the most plausible account of our options (or so I've argued) and, thus, the one that we should combine with maximalism. And when so combined, rationalist maximalism will be able to require agents to form certain attitudes and to perform certain mixed acts. For another, rationalist maximalism has the resources to successfully rebut objections to which other versions of maximalism fall victim. Perhaps, the most significant of these is the objection that maximalism has counterintuitive implications in cases that seem to have the following three features: (2.1) S has the options of φ -ing well, φ -ing poorly, and not φ -ing at all. (2.2) It would be okay if S doesn't φ at all, but it would be better (indeed, best) if she were to φ well. And worst of all would be if she were to φ poorly. Thus, she ought to φ well. (2.3) As a matter of fact, if S were to φ , she would φ poorly.

To illustrate, consider the now famous case of *Professor Procrastinate*:

Professor Procrastinate receives an invitation to review a book. He is the best person to do the review, has the time, and so on. The best thing that can happen is that he says yes, and then writes the review when the book arrives. However, suppose it is further the case that were Procrastinate to say yes, he would not in fact get around to writing the review. Not because of incapacity or outside interference or anything like that, but because he would keep on putting the task off. ...Moreover, we may suppose, [his saying yes and never writing the review] is the worst that can happen. It would lead to the book not being reviewed at all. (Jackson & Pargetter 1986, 235)

In this case, S is Professor Procrastinate, φ -ing is accepting the invitation, φ -ing well is accepting and then writing, and φ -ing poorly is accepting and then never writing.

Employing this sort of case, critics offer the following argument against maximalism:¹⁴

- (2.4) Maximalism is true, and, thus, *S*'s φ -ing is permissible if and only if it is entailed by some permissible option. [Assumption for *reductio*]
- (2.5) *S* has the option of φ -ing well. [Assumption]
- (2.6) If φ -ing well is an option, it is *S*'s best option. And, thus, if φ -ing well is an option, it's a permissible option (indeed, it is what *S* ought to do). [From the stipulations of the case]
- (2.7) φ -ing well entails φ -ing. [Analytic]
- (2.8) Thus, *S*'s φ -ing is permissible. [From 2.4–2.7]
- (2.9) *S*'s φ -ing is not permissible. [Intuition]
- (2.10) Therefore, it is not the case that maximalism is true. [From 2.4, 2.8, & 2.9]

Such critics find 2.9 intuitively obvious. They claim that Professor Procrastinate is not permitted to accept given that he would not write if he were to accept.¹⁵

Now, not everyone finds 2.9 intuitively compelling, but nevertheless that's the argument against maximalism that many critics give. I suspect that disagreement about 2.9 stems from the fact that such cases (e.g., *Professor Procrastinate*) are under-describe. And, once we take note of the two ways in which the missing details might be spelled out, we see that we should reject either 2.5 or 2.9. Here's my argument for this.

¹⁴ See, for instance, Cariani 2013, Jackson & Pargetter 1986, and Snedegar 2014. Admittedly, they don't use this sort of argument directly against maximalism. Instead, they use it against views like possibilism (i.e., the view that whether one ought to φ depends on whether the best possible thing that one can do entails φ -ing) and inheritance (i.e., the view that if φ -ing entails ψ -ing, then '*S* ought to φ ' entails '*S* ought to ψ '). But, given the relevant similarities between these views, the argument can be turned into one against maximalism, as I've done above.

¹⁵ It may seem that 2.9 just follows via *modus ponens* from both the fact that (F1) if Professor Procrastinate would not write if he were to accept, then he is not permitted to accept and the fact that (F2) Professor Procrastinate would not write if he were to accept. But it does so only if we assume that F1 is a material conditional with "Professor Procrastinate would not write if he were to accept" as its antecedent and "he is not permitted to accept" as its consequent. But many deny this assumption—see, for instance, those who think that deontic conditionals are a special, primitive type of conditional (e.g., von Wright 1956) and also those who think that the prohibition has wide-scope over the entire conditional (e.g., Broome 2004). In any case, we can clearly see that the inference to 2.9 from F1 and F2 is invalid: just consider the Paradox of Gentle Murder (Forrester 1984): (1) If I'm going to kill Mikhail, then I'm not permitted refrain from killing him gently. (2) As a matter of fact, I'm going to kill him. Therefore, (3) I'm not permitted to refrain from killing him gently. But assume that there is no good reason to kill Mikhail and many good reasons not to. Thus, assume that I'm obligated to refrain from killing him, gently or otherwise.

- (2.11) S is either irrepressible or repressible—that is, either (a) S will φ poorly regardless of how he now responds to his reasons or (b) it is not the case that S will φ poorly regardless of how he now responds to his reasons. [From the law of excluded middle]
- (2.12) If S is irrepressible, then we should reject 2.5, which says that S has the option of φ -ing well. For if S will φ poorly regardless how he now responds to his reasons, then he doesn't, at present, have rational control of whether she φ s well. And if she doesn't, at present, have rational control of whether she φ s well, she doesn't, at present, have the option of φ -ing well. [From rationalism]
- (2.13) If S is repressible, then we should reject 2.9, which says that S's φ -ing is not permissible. For, if he is repressible, then he should direct the course of his future actions by responding appropriately, at present, to his reasons, thereby ensuring that he will φ and φ well. And, thus, he's not just permitted to φ ; he's obligated to φ . [Intuition]
- (2.14) Therefore, we should reject either 2.5 or 2.9. [From 2.11–2.13]

To illustrate, consider *Professor Procrastinate*. Given Jackson and Pargetter's description of the case, it's unclear whether Professor Procrastinate is repressible or irrepressible. For all that they say, it could be that Procrastinate is aware of his tendency to procrastinate and that, when it's really important to him that he doesn't procrastinate, he resolves now not to give into the temptation to procrastinate later on. And it may even be, as we'll indeed suppose, that his making this resolution is sufficient to ensure that he won't procrastinate. And, in that case, he is repressible, for he will write the review if he responds appropriately to his reasons by resolving now to write the review as soon as the book arrives. So, one possibility for why he wouldn't write if he were to accept is that he's not now responding appropriately to his reasons. And, in that case, my clear intuition is that Professor Procrastinate is not only permitted to accept, but is obligated to accept. For he's obligated to respond appropriately to his reasons, accepting the invitation while also

resolving to write the review as soon as the book arrives. In which case, he will accept and write the review. And so it seems that we should reject 2.9 if he is irrepensible.¹⁶

Of course, he could instead be irrepensible such that no matter how he responds now to his reasons, and, thus, no matter what he resolves now to do later on, his later self is going to choose to procrastinate when the book arrives. In that case, I think that he just has to accept that he is irrepensible, having no more control at present over whether his future self will write the review than I have over whether the next U.S. Congress will amend the constitution so as to prohibit the private ownership of firearms. And, if that's how we're supposed to imagine things, then although we should readily accept that Professor Procrastinate is not permitted to accept (and, so, accept 2.9), we should deny that he has, at present, the option of accepting and writing. For if he doesn't, at present, have the power to direct the course of his future actions so as to ensure that he will write the review when the book arrives, in what sense does he have, at present, the option of writing when the book arrives? To have the option of ϕ -ing is to have control over whether one ϕ s. But, if Professor Procrastinate is irrepensible, he has, at present, no control over whether he will write when the book arrives. And so I think that we should deny 2.5 if Procrastinate is irrepensible.

Admittedly, if we were to accept schedulism, we would have to accept 2.5. For when the book arrives Procrastinate will have the capacity to form the intention to start work on the review right away. And he would, we'll assume, start work on the review as soon as the book arrived if he were to form that intention when the book arrived. But if he's irrepensible, then no matter how he responds to his reasons now (and no matter what he resolves now to do later on), he won't form the intention to start work when the book arrives. Instead, he'll form the intention to put it off until next week. And when that week arrives, he'll form the intention to put it off another week. And so and so forth. So if we were to adopt schedulist maximalism, we would have to accept that even an irrepensible

¹⁶ Another possibility is that he has made the following prior arrangement with a colleague: If he copies his colleague on an email in which he accepts an invitation to do something, he thereby bets this colleague ten thousand dollars that he will do that something. Further suppose that, given this, Professor Procrastinate would write the review if he were to copy his colleague on the email in which he accepts the journal's invitation to write the book review. And this is compatible with Jackson and Pargetter's description of the case, for it may be that Professor Procrastinate doesn't want to make this bet and so wouldn't copy his colleague on the email if he were to accept. And, in that case, it could still be that, were Professor Procrastinate to accept, he would not write the review. But, in such a case, I do not have the intuition that Professor Procrastinate should not accept the invitation (that is, I do not find 2.9 plausible). Rather, I have the intuition that Professor Procrastinate should accept by email while copying his colleague on that email and should, therefore, accept.

Procrastinate has the option of accepting and writing and, so, ought to accept and write—this being his best option. And, on maximalism, that entails that he ought to accept. And if Procrastinate is irrepressible, this does seem counterintuitive. But if we accept rationalist maximalism, we can just deny that an irrepressible Procrastinate has the option of accepting and writing. And, thus, we can accept the intuitive thought that an irrepressible Procrastinate ought not to accept the invitation. So whereas schedulist maximalism may fall victim to this objection, rationalist maximalism does not.

This isn't the only objection that maximalism can avoid by adopting rationalism. Consider Johan Gustafsson's recent objection to maximalism.¹⁷ He argues that maximalism has counterintuitive implications in the following sort of case.

Newcomb's Non-Problem: At t_0 , you are offered a chance to participate in a version of Newcomb's problem, which involves two boxes: one transparent and one opaque. At t_1 , you must choose whether to participate. If you agree at t_1 to participate, you must take at t_3 either just the transparent box or both boxes. And, as you are told, you must take possession of the contents of whatever box or boxes you take, whether they be good or bad. Thus, if you take both boxes, and the opaque box happens to contain a writ of debt for a \$1,000,000, you will be responsible for paying that debt. Now, as you can see, the transparent box contains \$1,000. But the contents of the opaque box are a complete mystery to you—that is, until t_2 , which is when you are given the following additional information: the opaque box contains either \$1 or \$1,000,001. It contains \$1 if you formed at t_1 the intention to take both boxes. Otherwise, it contains \$1,000,001.

Even if you end up choosing at t_1 to participate and take at t_3 both boxes, there are two possibilities, which have very different outcomes:

(Poss1) At t_1 , you chose to participate. At t_3 , you took both boxes. But you did not form the intention to take both boxes until t_2 , which is when you

¹⁷ See his 2014.

learned that there's nothing bad in the opaque box. Consequently, you ended up with \$1,001,001.

(Poss2) At t_1 , you chose to participate. And, at t_1 , you formed the intention to take both boxes at t_3 even though you did not at the time have any idea what was in the opaque box. At t_3 , you took both boxes. You ended up with only \$1,001.

Clearly, if Poss2 is actualized, you have failed in some way. For, as Gustafsson notes, it was up to you at t_0 which of Poss1 and Poss2 would be actualized. And if Poss2 is actualized, then you ended up with a million fewer dollars. Of course, from the fact that you ended up with less money, it doesn't necessarily follow that you failed in some way. But the thought is that you ended up with less money than you should have ended up with. For you should not have formed at t_1 the intention to take both boxes at t_3 . That was reckless given that, for all you knew, the opaque box contained a writ of debt for a \$1,000,000 or more. Moreover, had you not recklessly formed this intention at t_1 , you would have actualized Poss1, ending up with an additional million dollars. Thus, as Gustafsson claims, it seems that a plausible theory would require you to actualize Poss1 as opposed to Poss2. The problem with maximalism, Gustafsson believes, is that maximalism cannot do this. But while it's true that schedulist maximalism cannot do this, rationalist maximalism can.

On schedulist maximalism, your only options are (O1) choose at t_1 to participate and take at t_3 both boxes, (O2) choose at t_1 to participate and take at t_3 only the transparent box, and (O3) choose at t_1 not to participate. These are the only options, because, on schedulism, only that which can be done intentionally counts as an option, and your forming or refraining from forming an intention is not something that you can do intentionally—that is, it's not something that you can do by intending to form (or intending not to form) that intention.¹⁸ So the best schedulist maximalism can do is require you to perform O1, but your

¹⁸ To see that the formation of an intention is not something one can do at will, consider the toxin puzzle (Kavka 1983). I will receive a million dollars tomorrow morning if and only if, at midnight tonight, I intend to drink some toxin tomorrow afternoon. Drinking the toxin will not kill me, but it will make me terribly ill for several days. Whether I receive the million dollars tomorrow morning depends only on what I intend to do at midnight tonight, not on whether I drink the toxin tomorrow afternoon. Realizing this, I'm unable to intend at midnight tonight to drink the toxin tomorrow afternoon. For I see no reason to drink the toxin. I know that, come tomorrow afternoon, I'll either have the million dollars or I won't. And in

performing O1 doesn't necessitate your actualizing Poss1 as opposed to Poss2. Thus, there is no way for schedulist maximalism to require you to actualize Poss1 as opposed to Poss2.

Fortunately, rationalist maximalism, unlike schedulist maximalism, doesn't fall victim to this objection. On rationalist maximalism the relevant options are (O2) choose at t_1 to participate and take at t_3 only the transparent box, (O3) choose at t_1 not to participate, (O4) choose at t_1 to participate while simultaneously forming the intention to take both boxes and then take at t_3 both boxes, and (O5) choose at t_1 to participate while not forming at t_1 the intention to take both boxes, form at t_2 the intention to take both boxes, and take at t_3 both boxes. Since rationalist maximalism allows that O5 is an option, and since your performing O5 necessitates your actualizing Poss1, rationalist maximalism can require you to actualize Poss1 by requiring you to perform O5. Thus, rationalist maximalism can avoid Gustafsson's objection.¹⁹

So, we've seen that if we combine maximalism with rationalism, maximalism avoids the sorts of objections that have typically been raised against it.²⁰ But rationalist maximalism is not just the most plausible version of maximalism, it is plausible in its own right, for, as I'll show in the next section, it is uniquely well-suited to accommodate our intuition that a moral theory should be collectively successful.

3. The Principle of Moral Harmony

To understand what it means for a theory to be *collectively successful*, let's start by looking at a theory that isn't: viz., rational egoism.²¹ It's the view that S's φ -ing is rationally permissible

neither case will I have any reason to drink toxin. So, I have decisive reason not to drink the toxin. Given this, I'm unable to form the intention to drink the toxin. And this shows that forming an intention is not typically something that I can do at will—that is, voluntarily.

¹⁹ See Bykvist 2002 for another objection to maximalism, and see Gustafsson 2014 (587–588) for some discussion. As with the objections discussed in the body of this paper, it relies on an implausible account of what our options are. For Krister Bykvist's objection to even get off the ground, he must presume that S has, as of t , the option of φ -ing at t' only if she can form at t the intention to φ at t' . On this view, I don't now have the option of forming tomorrow the belief that there are two sticks of butter in my fridge (assume that tomorrow there will be exactly two sticks of butter in my fridge) even though I can now form the intention to look inside my fridge tomorrow and would form this belief if I were to do so. On Bykvist's view, I don't now have the option of forming this belief tomorrow, because I cannot now form the intention to form this belief. For I'm a rational person, and a rational person cannot form the belief to φ when she knows that φ is not something that she can intentionally do.

²⁰ There are other objections to maximalism that, unlike the ones discussed in the body of this paper, don't depend on our accepting an implausible account of what our options are. For instance, some object to maximalism on the grounds that it gives rise to Ross's paradox, and it does so regardless of what account of options we accept. I address this objection in another paper—see my 2015.

²¹ I borrow the term 'collectively successful' from Parfit 1984 (92).

if and only if φ -ing is in S 's best interests. And it gives each of us the aim of promoting his or her own interests. This theory is not collectively successful because having everyone satisfy its requirements doesn't guarantee that our theory-given aims will be best achieved. In fact, the universal satisfaction of rational egoism can actually be self-defeating. For, given the prevalence of collective-action problems, our theory-given aims will often be better achieved if none of us satisfy its requirements. To illustrate, consider *The Prisoner's Dilemma*:

Two members of a criminal-gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of communicating with the other. The prosecutors lack sufficient evidence to convict the pair on the principal charge. They hope to get both sentenced to a year in prison on a lesser charge. Simultaneously, the prosecutors offer each prisoner a bargain. Each prisoner is given the opportunity either to: betray the other by testifying that the other committed the crime, or to cooperate with the other by remaining silent. The offer is:

- If A and B each betray the other, each of them serves 2 years in prison
- If A betrays B but B remains silent, A will be set free and B will serve 3 years in prison (and vice versa)
- If A and B both remain silent, both of them will only serve 1 year in prison (on the lesser charge). (Wikipedia contributors 2015)

It's in the best interests of each prisoner to betray the other. To see this, imagine that you are one of the prisoners. On the one hand, if you are going to be betrayed by the other prisoner, then you are better off betraying her, ensuring that you get two years instead of three years. And if, on the other hand, you aren't going to be betrayed by her, then you are still better off betraying her, ensuring that you get zero years instead of one year. So, either way, you are better off betraying her. And the same reasoning applies for her. So she's better off betraying you. Thus, if you both satisfy the requirements of rational egoism, you will each betray the other, ending up with two years in prison apiece. But if you had both violated rational egoism's requirement and remained silent, you would have each ended up with only one year in prison. Thus, rational egoism is self-defeating. In *The Prisoner's Dilemma*, the prisoners will better achieve their theory-given aim of promoting their own interests if they both violate as oppose to abide by its requirements.

This, of course, doesn't mean that rational egoism is false. Rational egoism is a theory about whether an individual's actions are rationally permissible, not a theory about whether a group's collective actions are rationally permissible, nor a theory about whether an individual's actions are morally permissible. So although we should clearly expect rational egoism to be *individually successful* such that no individual will ever better achieve her theory-given aims by not satisfying it, we shouldn't necessarily expect it to be *collectively successful* such that everyone's theory-given aims will be better achieved by everyone's satisfying it.

But many believe that even if a theory of individual rationality needn't be collectively successful, a theory of morality must.²² Morality must, it seems, be something that it would be good for us as a group to follow. Thus, the universal satisfaction of morality must not be self-defeating. Quite the opposite. The universal satisfaction of morality by all agents at all times must guarantee that we end up in the morally best world that we could reasonably be expected to bring about. To accept that a moral theory must be collectively successful in this sense is to accept what's sometimes called the *principle of moral harmony* (Feldman 1980). Exactly how this principle should be formulated is controversial, but its gist should be clear enough by now. In any case, my plan is to provide some further motivation for the principle before looking at precisely how it should be formulated. So consider the following example, which I borrow, with some modification, from David Estlund (Forthcoming).

Slice and Patch Go Golfing: Unless a patient's tumor is removed very soon, she'll die (though not painfully). Immediate surgery and stitching by the only two available doctors, Dr. Slice and Dr. Patch, is the one thing that will save her life. But if there is surgery without stitching, her death will be agonizing. And if there is stitching without surgery, her death will likewise be agonizing. It would even be cruel for one of them to show up to the hospital knowing that the other won't, for this would only needlessly get the patient's hopes up, making her death psychologically agonizing. So, it seems that Dr. Slice, who doesn't know how to stitch, ought to show up to the

²² See, for instance, Baier 1958, Castañeda 1974, Parfit 1984 (94), Pinkert 2015, Regan 1980, and Zimmerman 1996. For more examples, see Feldman 1980. For arguments against the view that morality must be collectively successful, see Feldman 1980 and Kierland 2006.

hospital to perform the surgery if and only if Dr. Patch will be there to stitch her up afterwards. And Dr. Patch, who doesn't know how to perform the surgery, ought to be there to give her stitches if and only if Dr. Slice will be there to perform the surgery. Unfortunately, Slice and Patch are each going golfing whether the other attends to the patient or not, because neither cares at all whether the patient lives or dies. And each knows this about the other. Predictably, then, the patient dies (though not painfully) while Slice and Patch enjoy a pleasant round of golf.

As Estlund (Forthcoming) points out, "many of us respond to this case with the intuition that there is some moral violation here, but the puzzle is to find an agent who has committed it." It seems that Dr. Slice was under no obligation to be at the hospital to perform the surgery given that, as he knew, no one was going to be there to stitch the patient afterwards. And it seems that Dr. Patch was under no obligation to be at the hospital to stitch the patient given that, as he knew, no one was going to be there to perform the surgery. Still, it seems that there must have been some moral violation. After all, the group consisting of Slice and Patch could have brought about a significantly better world: the one in which the patient lives. So how could it be that they each did all that they were morally required to do and yet they failed to bring about the morally best world that the two of them were capable of bringing about? Of course, to even pose this rhetorical question is to presuppose that the satisfaction of a moral theory by some group must guarantee that they end up in the morally best world that they are capable of bringing about. Thus, it seems that if we are to accommodate the intuition that there has been a moral failure in this case, we must accept the principle of moral harmony – that is:

(PMH) A moral theory, T, is correct if and only if the agents who satisfy T, whoever and however numerous they may be, are guaranteed to produce the morally best world that that they could together bring about.²³

²³ This is based on Donald Regan's definition of 'adaptability': "a theory T is adaptable if and only if the agents who satisfy T, whoever and however numerous they may be, are guaranteed to produce the best consequences possible as a group, given the behaviour of everyone else" (1980, 6). The difference is that whereas Regan is concerned only with consequentialist moral theories, I'm concerned with moral theories in general. Thus, the morally best world could, if, say, Kantianism is correct, be the

However plausible the idea that a moral theory must be collectively successful is, PMH is too strict a requirement. For, if PMH were true, no moral theory could be correct. Donald Regan proved this back in 1980 with his excellent book *Utilitarianism and Co-operation*.²⁴ Let me explain his reasoning. First, Regan shows that no moral theory that is exclusively act-orientated can satisfy PMH, where a theory is exclusively act-orientated if and only if it requires only that agents perform and refrain from performing certain voluntary acts.²⁵ To illustrate, consider that, in *Slice and Patch Go Golfing*, neither Slice nor Patch perform any immoral act. After all, there is nothing immoral about going golfing when there is nothing better that one could be doing. And, given that each is unwilling to attend to the patient and is going golfing regardless of what the other does, neither can do anything to save the patient's life. Each can only make things worse for the patient by showing up to the hospital, giving her false hope and, thereby, making her death psychologically agonizing. Thus, if we are to claim that there has been a moral violation in this case, the violation must lie, not with their voluntary actions, but with something else. In this case, the violation seems to lie with neither of them caring whether the patient lives or dies and/or with each of them intending to go golfing regardless of what the other does. But to refrain from caring whether the patient lives or dies is not to refrain from performing a voluntary act. Nor is intending to go golfing regardless of what the other does a voluntary act. Thus, the only way for a moral theory to pass PMH is for it to require that agents not only perform certain voluntary acts, but also form (non-voluntarily) certain attitudes, such as the desire the patient lives and/or the intention to attend to the patient's needs should the other also be willing to do so.

Second, Regan shows that any moral theory that is not exclusively act-orientated will violate PMH. For any moral theory that's not exclusively act-oriented will have to require 'something more' of agents than just the performance (or non-performance) of certain voluntary acts. And, as Regan notes, "there is always the possibility that there will be a mad

world in which the group commits the fewest and/or least significant violations of the categorical imperative possible as opposed to the world in which the group produces the best consequences possible.

²⁴ He proved that no theory can be adaptable unless we ignore the direct consequences of applying that theory's decision procedure. See his 1980, chapter 10.

²⁵ This is not Regan's definition, for he provides no definition—see 1980, 109. But I believe that this definition captures (at least, sufficiently well for our purposes) the notion that he has in mind.

telepath...who will blow up Macy's [or the whole planet] in response to that 'something more'" (1980, 181). And, of course, a world in which the whole planet is blown up is not going to be the morally best world that Slice and Patch could bring about. For they could just refrain from this something more, in which case the planet would be spared. And it's better that one patient dies painlessly than that everyone on the planet, including the patient, dies in a horrible explosion. So consider a revised version of *Slice and Patch Go Golfing*, which I'll call *The Mad Telepath*. In this case, everything is the same as in original but for the addition of a mad telepath who will blow up the whole planet if either Slice or Patch form the desire that the patient lives (or even if either of them forms the intention to attend to the patient's needs should the other also be willing to do so). So, unlike *Slice and Patch Go Golfing*, this is a case where Slice and Patch would not bring about the morally best world by desiring that the patient lives and intending to attend to the patient's needs so long as the other is also willing to do so.

Together, these two points imply that there is no way for a moral theory to satisfy PMH. A moral theory will either have to be exclusively act-orientated or not. If, on the one hand, it is exclusively act-orientated, then there will be instances, such as in *Slice and Patch Go Golfing*, where a group of agents who all satisfy the theory fail to bring about the morally best world that that they can bring about, because they fail to have certain attitudes. If, on the other hand, a moral theory is not exclusively act-orientated, then there will be instances, such as in *The Mad Telepath*, where a group of agents who all satisfy the theory fail to bring about the morally best world that that they could bring about, because they have those same attitudes and a mad telepath is going to destroy our planet because of this. So either way there will instances in which the moral theory is satisfied by everyone in the group and yet the group fails to bring about the morally best world that they could bring about.

Clearly, then, PMH is too strict. I think that PMH goes wrong in insisting that a moral theory must be such that if, in *The Mad Telepath*, Slice forms the desire that the patient lives, he must have thereby violated a moral requirement, since his forming this desire results in the production of a suboptimal world—specifically, the one in which the mad telepath destroys the planet. This, I believe, is a mistake, because Slice doesn't have voluntary control over whether he refrains from forming this desire. And this means that he cannot refrain from forming this desire for whatever reason he takes to be sufficient reason

to do so. So, even if he takes the fact that his forming this desire would lead to the destruction of the planet as sufficient reason for him to refrain from forming this desire, he cannot refrain from forming this desire for this reason. Indeed, given his lack of voluntary control, he can refrain from forming this desire only for the sort of reason that would make it fitting for him to so refrain, such as the reason that he would have to so refrain if the patient's continued life were undesirable. But the patient's continued life *is* desirable. Thus, the only way that he can refrain from forming this desire is by failing to respond appropriately to the decisive reason that he has to form this desire — i.e., the fact that the patient's continued life is desirable. And I don't think that morality can reasonably require us to fail to respond appropriately to our reasons just because our so failing would prevent some disaster. That is, I don't think that morality can require us either to form attitudes for which we lack sufficient reason or to refrain from forming attitudes for which we have decisive reason.²⁶

To see why, consider the following example.

Hating and Saving Rocks: Unless I both hate professional wrestler Dwayne Johnson (a.k.a. The Rock) because of his Samoan ancestry and intend to kill him with my bare hands in a fair fight, an evil demon will destroy the third rock from the sun (that is, our planet).

Note that the fact that Dwayne Johnson is of Samoan ancestry is no reason to hate him. And I can't just voluntarily hate him for this reason. I cannot even hate him for the reason that hating him might save the planet. Given that I don't have voluntary control over whether I hate him, I cannot hate him for whatever reason I take to be sufficient for hating him. I can hate him only for the reason that I think him despicable or otherwise deserving of hatred. The problem is that I don't think that he is despicable. Certainly, neither the fact that he is of Samoan ancestry nor the fact that an evil demon will destroy our planet if I don't hate him make him despicable. So I can't hate him — at least, not insofar as I respond appropriately to my reasons. What's more, I can't form the intention to kill him with my bare hands in a fair

²⁶ As I'll use the terms, S has *decisive reason* to φ if and only if S's reasons are such as to make S obligated to φ , and S has *sufficient reason* to φ if and only if S's reasons are such as to make S permitted to φ .

fight insofar as I respond appropriately to my reasons. For I know that I cannot take him in a fair fight given his massive physique and superior fighting abilities. Moreover, even if I could, I would not intend to kill him, for I have no good reason to kill him and many good reasons not to. And so, if I respond appropriately to my reasons, I will not intend to kill him with my bare hands in a fair fight. Thus, responding appropriately to my reasons precludes me from forming the attitudes that I must form in order to prevent the destruction of our planet.

Now, it would be very strange to think that morality could require me to respond inappropriately to my reasons given that what makes me the sort of subject to which moral obligations and responsibilities apply is that I'm the sort of subject who's capable of responding appropriately to my reasons—a rational agent. It seems nonsensical for some moral requirement to apply to me because I have the capacity to respond appropriately to my reasons when I can fulfill that requirement only by failing to respond appropriately to my reasons.²⁷ Such would be the nature of a moral requirement for me either to hate Johnson because of his Samoan ancestry or to intend to kill him with my bare hands in a fair fight. Thus, I think that it is a mistake to think, as PMH supposes, that morality can require us to respond inappropriately to our reasons by either forming attitudes for which we lack sufficient reason (e.g., my forming a hatred for Johnson) or refraining from forming attitudes for which we have decisive reason (e.g., Slice's refraining from forming the desire that the patient lives).

If this is right, we must revise PMH as follows.

(PMH*) A moral theory, *T*, is correct if and only if the agents who satisfy *T* while otherwise responding appropriately to their reasons, whoever and however numerous they may be, are guaranteed to produce the morally best world that is compatible with each of them responding appropriately to their reasons.²⁸

²⁷ For more on this, see Portmore 2011, 38–51.

²⁸ My formulation of the principle of moral harmony is distinct from those that have been proposed by Pinkert (2015), Regan (1980), and Zimmerman (1996). I don't have space here to explain why I reject each of these alternative proposals, but see Kierland 2006 and Forcehimes & Samrau 2015 for criticisms of each of these proposals. Now, Kierland would reject my proposal as well on the grounds that it violates the principle that 'ought' implies 'can'. But it violates this principle only if we assume that the relevant sense of 'can' is to be analyzed in terms of voluntary control as opposed to rational control, as I've

PMH* allows us to account for our intuition that there must have been some moral failure in *Slice and Patch Go Golfing*. The moral failure lies with the fact that Slice and Patch failed to have the desires and intentions that they were morally required to have. They were morally required to desire that the patient lives and to intend to attend to her so long as the other is also so willing. And if they both had these attitudes, they would have each been such that the other was required to attend to the patient. Thus, abiding by a moral theory that passes PMH* ensures that Slice and Patch produce, in *Slice and Patch Go Golfing*, the morally best world that they are capable of producing: the one in which the patient lives. But PMH* does not imply that abiding by the correct moral theory ensures that Slice and Patch prevent, in *The Mad Telepath*, the destruction of our planet. Abiding by a moral theory that passes PMH* ensures only that Slice and Patch produce the morally best world that's compatible with their responding appropriately to their reasons—that is, the morally best world that they can reasonably be expected to bring about. And for them to respond appropriately to their reasons, they must, or so I'm assuming, form the desire that the patient lives and form the intention to attend to her so long as the other is also willing to do so. So, unfortunately, the morally best world that's compatible with their responding appropriately to their reasons is, in *The Mad Telepath*, the world in which the mad telepath destroys our planet. But this is no reason to reject PMH*. After all, we should not expect the correct moral theory to ensure that no disaster ever befalls us. For no matter what attitudes a moral theory requires of us, a mad telepath can always wreak havoc upon us for fulfilling those requirements. Moreover, we should think that the correct moral theory will never require us to respond inappropriately to our reasons given that our being rational is what subjects us to moral obligations and responsibilities in the first place.

Now, according to PMH*, a moral theory, T, will be correct if and only if the agents who satisfy it while otherwise responding appropriately to their reasons, whoever and however numerous they may be, are guaranteed to produce the morally best world that is compatible with each of them responding appropriately to their reasons. And the only sort

argued. And Forcehimes & Samrau 2015 would probably question the need to appeal to the principle of moral harmony at all and suggest that we would do better to locate Slice's and Patch's moral failures with whatever past acts that they performed that led them to have their problematic attitudes. But see my example involving Jane and her belief that the Earth is no more than a few thousand years old as well as Smith 2015 for why I think this sort of tracing strategy won't work.

of theory that can do that is one that, like rationalist maximalism, holds that the permissibility of some particular act depends on whether it is part of some permissible whole that includes both the performance of various voluntary acts and the formation of various reasons-responsive attitudes. Indeed, the only sort of theory that I can see satisfying PMH* is some version of rationalist maximalism that takes 'F' in Max2 to stand for something like the following: "includes all and only those attitudes that are fitting and includes all and only those acts that would, if the agent were to have all and only fitting attitudes, actualize the morally best world that could be actualized by any of the acts available to the agent."²⁹ Thus, it seems that rationalist maximalism is uniquely well-suited to accommodate the idea that the correct moral theory must be collectively successful, where this is interpreted as satisfying PMH* as opposed to PMH.³⁰ Thus, insofar as we think that a moral theory must be collectively successful in this sense, we have reason to think that rationalist maximalism (or something very much like it) is the correct moral theory.

4. Conclusion

Rationalist maximalism is, I've argued, the most plausible version of maximalism. For, as I've argued, rationalism is the most plausible account of our options, and so the one that we should combine with maximalism. Moreover, I've argued that when we do combine

²⁹ If we interpret rational maximalism as a theory about *moral* permissibility and make this substitution for F, rational maximalism will entail that a maximal option is morally permissible only if it includes all those attitudes that are fitting. This may seem problematic since, in many instances, the fittingness of an attitude will have nothing to do with morality. For instance, the fitting response to the perception of a bright flash of lightning is often the belief that one will hear a loud clap of thunder very shortly. Nevertheless, there wouldn't necessarily be anything morally impermissible about a maximal option that failed to include this fitting attitude. I suggest, then, that when we make this substitution for F, we also modify Max2 so that we get: "(Max2) for any maximal option μ , S 's μ -ing is both rationally and morally permissible if and only if S 's μ -ing includes all and only those attitudes that are fitting and includes all and only those acts that would, if the agent were to have all and only fitting attitudes, actualize the morally best world that could be actualized by any of the acts available to the agent, and when S 's μ -ing is both rationally and morally permissible, this is in virtue of the fact...."

³⁰ Rationalist maximalism also has the benefit of avoiding Feldman's objection to PMH. Feldman asks us to imagine "that, a group of adults has taken a group of children out to do some ice skating. The adults have assured the children and their parents that, in case of accident, they will do everything in their power to protect the children. ...A lone child is skating in the middle, equidistant from the adults. Suddenly, the ice breaks, and the child falls through. There is no time for consultation or deliberation. Someone must quickly save the child. However, since the ice is very thin, it would be disastrous for more than one of the adults to venture near the place where the child broke through. For if two or more were to go out, they would all fall in and all would be in profound trouble. In fact, let us suppose, no one goes to the aid of the child" (1980, 171–172). Feldman believes that each adult was morally obligated to quickly head out to the hole in the ice to rescue the child. But, of course, if each adult did this, disaster would ensue. Thus, Feldman concludes that PMH is false. But if we accept rationalist maximalism, we should say that the relevant moral obligation was not merely to head out to the hole in the ice to rescue the child, but to do so while intending to stop abruptly should there be any indication that others are heading out toward the hole as well. And it is not the case that if each adult satisfied this moral requirement, disaster would ensue.

rationalism with maximalism, the result is a theory that can avoid the sorts of objections that have typically been levelled against maximalism. And I've done more than just argue that rationalist maximalism is the most plausible version of a certain class of moral theories (viz., maximalist theories). I've also argued that, insofar as we find PMH* plausible, we have reason to think that it is the correct moral theory.

Of course, even if rationalist maximalism is the correct moral theory, that doesn't mean that we must reject wholesale all alternative moral theories. For rationalist maximalism doesn't provide a substantive account of what we should aim at, morally speaking. Should we, for instance, aim solely at making the world as good as possible or also at respecting people's autonomy even when doing so will make the world worse? Thus, even if we should, as I propose, adopt a version of rationalist maximalism according to which 'F' in Max2 stands for something like "includes all and only those attitudes that are fitting and includes all and only those acts that would, if the agent were to have all and only fitting attitudes, actualize the morally best world that could be actualized by any of the acts available to the agent," we still need to know which of the various possible worlds that the agent could actualize is the morally best one from her perspective. Is it, for instance, the one where she commits one murder so as to prevent five others from each committing one comparable murder or is it the one in which there are more murders but none that have been committed by her? Theories like Kantianism and utilitarianism can help us answer such questions. Still, even if we are not to reject such theories wholesale, we must, if rationalist maximalism is correct, reject them in their current forms. As things stand, such theories assess only acts and assess each act (whether maximal or non-maximal) according to the same standards. But if rationalist maximalism is correct, this is a mistake. If we are to accept rationalist maximalism, we should think that moral theories must require us to form of certain reasons-responsive attitudes, and not just to perform certain voluntary acts. What's more, they must evaluate each non-maximal act in terms of whether its performance is entailed by that of some permissible maximal option, not in terms of whether it maximizes utility or accords with Kant's categorical imperative.³¹

³¹ For helpful comments and discussions on precursors to this paper, I thank audiences at Syracuse University, Arizona State University, Australian National University, University of Colorado, Boulder, University of Maryland, College Park, the 2014 Rocky Mountain Ethics Congress, the Arizona Center for the Philosophy of Freedom, and the 2014 Pacific Division Meeting of the American Philosophical Association. I especially thank Chrisoula Andreou, David Boonin, Cheshire Calhoun, Eric Chwang, Stew Cohen, Yishai Cohen, Brad Cokelet, Dale Dorsey, Josh Gert, Peter A. Graham, Pat Greenspan, Johan E. Gustafsson, Chris

Arizona State University
dwpportmore@gmail.com

References

- Baier, K. (1958). *The Moral Point of View*. Ithaca, NY: Cornell University Press.
- Broome, J. (2004). "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. 28–55. Oxford: Oxford University Press.
- Bykvist, K. (2002). "Alternative Actions and the Spirit of Consequentialism." *Philosophical Studies* 107: 45–68.
- Cariani, F. (2013). "'Ought' and Resolution Semantics." *Noûs* 47: 534–558.
- Castañeda, H.-N. (1974). *The Structure of Morality*. (Springfield, Ill.: Charles Thomas Publisher).
- Estlund, D. (Forthcoming). "Prime Justice." In K. Vallier and M. Weber (eds.), *Political Utopias*. Oxford: Oxford University Press.
- Feldman, F. (1986). *Doing the Best We Can: An Essay in Informal Deontic Logic*. Dordrecht: D. Reidel Publishing Company.
- . (1980). "The Principle of Moral Harmony." *The Journal of Philosophy* 77: 166-179.
- Fischer, J. M., and M. Ravizza. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Forcehimes, A. T. & Semrau, L. (2015). "The Difference We Make: A Reply to Pinkert." *Journal of Ethics & Social Philosophy*, www.jesp.org, (September Discussion Note).
- Forrester, J. W. (1984). "Gentle Murder, or the Adverbial Samaritan." *Journal of Philosophy* 81: 193–196.
- Graham, A. (2014). "A Sketch of a Theory of Moral Blameworthiness." *Philosophy and Phenomenological Research* 88: 388–409.
- Goldman, H. S. [now H. M. Smith]. (1978). "Doing the Best One Can." In A. I. Goldman and J. Kim (eds.), *Values and Morals*, pp. 185–214. Dordrecht: D. Reidel Publishing Company.

Heathwood, Brian Hedden, Jonathan Herington, Adam Hosein, Victor Kumar, Frank Jackson, Eden Lin, Gene Mills, Dan Moller, Christopher Morris, G. Shyam Nair, Howard Nye, Graham Oddie, Andrew Reisner, Steve Reynolds, Melinda Roberts, Dan C. Shahaar, Walter Sinnott-Armstrong, Steve Sverdlik, Christian Tarsney, Larry Temkin, Travis Timmerman, Michael Tooley, Jean-Paul Vessel, Steve Wall, Ralph Wedgwood, Stephen White, and several anonymous referees. Work on this paper was supported by the RSSH Visiting Fellows Program, School of Philosophy, Australian National University.

- Gustafsson, J. E. (2014). "Combinative Consequentialism and the Problem of Act Versions." *Philosophical Studies* 167: 585–596.
- Hieronymi, (2008). "Responsibility for Believing." *Synthese* 161: 357–373.
- . (2006). "Controlling Attitudes." *Pacific Philosophical Quarterly* 87: 45–74.
- Jackson, F. and R. Pargetter (1986). "Oughts, Actions, and Actualism." *Philosophical Review* 95: 233–255.
- Kavka, G. (1983). "The Toxin Puzzle." *Analysis* 43: 33–36.
- Kierland, B. (2006). "Cooperation, 'Ought Morally', and Principles of Moral Harmony." *Philosophical Studies* 128: 381–407.
- King, A. (2014). "Actions that We Ought, But Can't." *Ratio* 27: 316–27.
- Maier, J. (2014). "Abilities." In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall Edition), URL = <<http://plato.stanford.edu/archives/fall2014/entries/abilities/>>.
- McHugh, C. (Forthcoming). "Attitudinal Control." *Synthese*.
- . (2014). "Exercising Doxastic Freedom." *Philosophy and Phenomenological Research* 88: 1–37.
- . (2012). "Epistemic Deontology and Voluntariness." *Erkenntnis* 77: 65–94.
- Parfit, D. (2011). *On What Matters*. Vol. 2. Oxford: Oxford University Press.
- . (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Pinkert, F. (2015). "What If I Cannot Make a Difference (and Know It)." *Ethics* 125: 971–998.
- Portmore, D. W. (Forthcoming). "Consequentialism and Coordination: How Consequentialism Has an Attitude Problem." In C. Seidel (ed.), *Consequentialism: New Directions, New Problems?* Oxford: Oxford University Press.
- . (2015). "Morality, Rationality, and Performance Entailment." Working manuscript that available at <http://bit.ly/1L6iVOI>.
- . (2013). "Perform Your Best Option." *The Journal of Philosophy* 110: 436–459.
- . (2011). *Commonsense Consequentialism: Wherein morality meets rationality*. New York: Oxford University Press.
- Regan, D. (1980). *Utilitarianism and Co-operation*. New York: Oxford University Press.

- Ross, J. (2012). "Actualism, Possibilism, and Beyond." In M. Timmons (ed.) *Oxford Studies in Normative Ethics: Volume 2*, pp. 74–96, Oxford: Oxford University Press.
- Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge, Mass.: Belknap Press.
- Smith, A. M. (2015). "Attitudes, Tracing, and Control," *Journal of Applied Philosophy*. Advance online publication available at <http://dx.doi.org/10.1111/japp.12107>.
- . (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115: 236–271.
- Snedegar, J. (2014). "Deontic Reasoning Across Contexts." In Cariani, F., Grossi, D., Meheus, J., and Parent, X. (eds.), *Deontic Logic and Normative Systems*, pp. 208–223. Switzerland: Springer International Publishing.
- von Wright, G. H. (1956). "A Note on Deontic Logic and Derived Obligation." *Mind* 65: 507–509.
- Wikipedia contributors. (2015). "Prisoner's dilemma." In *Wikipedia, The Free Encyclopedia*, https://en.wikipedia.org/w/index.php?title=Prisoner%27s_dilemma&oldid=681300872 (accessed September 21).
- Wiland, E. (2005). "Monkeys, Typewriters, and Objective Consequentialism." *Ratio* 18: 252–360.
- Zimmerman, M. J. (1996). *The Concept of Moral Obligation*. Cambridge: Cambridge University Press.